2001

Finite mixture models of heterogeneous capture probabilities for mark-recapture estimation of closed population size

Michael John Wallendorf
Iowa State University

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Finite mixture models of heterogeneous capture probabilities for mark-recapture estimation of closed population size

by

Michael John Wallendorf

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-majors: Statistics; Animal Ecology

Major Professors: Mark S. Kaiser and William R. Clark

Iowa State University

Ames, Iowa

2001

UMI Number: 3003275

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of

Michael John Wallendorf

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Co-major Program

Signature was redacted for privacy.

For the Co-major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

ABSTRACT

Heterogeneity in capture probabilities among animals is a common problem for estimation of animal population size from mark-recapture data. We model animals as belonging to discrete groups in which animals have the same probabilities of first capture and recapture. For removal data, population size, probabilities of first capture, and mixture proportions are estimated by maximum likelihood for a geometric finite mixture model. For mark-recapture data, a binomial finite mixture for recaptures is combined with a geometric finite mixture for first captures to better estimate mixture proportions. This model can be resticted for the assumption of no behavioral response to first capture.

On Carother's (1973) taxi cab data, estimation with a 2-group mark-recapture finte mixture provided a population size estimate , $N = 420$, that exactly matched the registered number of cabs. On meadow vole data, estimation with a 3-group model showed heterogeneity in behavioral response to first capture. Simulations show that our 2-group mark-recapture finite mixture estimator with restriction for no behavioral response to first capture is more efficient than Burham and Overtons' (1978) jackknife estimator when the smallest probability of first capture is 0.1 and the number of sampling occasions is 10.

CHAPTER 1. INTRODUCTION

For study of closed population size (N), the oldest mark-recapture method is the 2-event Petersen design (Seber 1982; Pollock 2000; Buckland, Goudie, and Borchers 2000). On occasion (event) one, n_1 animals are captured, marked, released, and allowed to mix with unmarked animals. On the second occasion, n_2 animals are captured or sighted and examined for marks. Among animals in the second occasion, m_2 marked animals are observed (recaptured). The numbers of animals caught in each occasion can be set by the amount of sampling effort the researcher applies on each occasion. The number of recaptures is the random variable for this process and in many studies (e.g., fish in lakes) it follows a hypergeometric distribution. Estimation of population size from these data requires the following assumptions 1) the population is closed (no addition and no deletion of animals during the study); 2) capture and marking in the first occasion does not affect the probability that a marked animal will be caught in the second occasion; and 3) the probability of capture is equal among animals in at least one of the two occasions. The maximum likelihood estimator of N from this model is known as the Lincoln-Petersen estimator, $\widehat{N} = n_1 n_2 / m_2$.

A Schnabel census is an extension of the Petersen design to allow the number of occasions, k to be greater than two. Animals in each sample are examined for marks (except occasion 1) and unmarked animals are given a unique mark that allows identification of the individual. After each occasion all animals are released to freely mix in the population. Each animal has a capture history corresponding to its status of capture on each occasion. Data from this design consist of a set of capture histories for animals caught during the study. This more extensive sampling enables the modeling of changes in capture probability through time, behavioral response to first capture, and heterogeneity of capture probabilities among animals. Otis et

al. (1978) consider a set of eight multinomial models based on the combinations of these three factors, time, behavior, and heterogeneity.

During a study, changes in environmental conditions (e.g., rainy weather, cool temperatures) can affect capture probabilities through time. A model for this effect allows capture probabilities to vary among occasions, but requires assumptions that the capture probabilities are constant among animals and have no change resulting from response to first capture. These models are widely applied to terrestrial studies particularly small mammals. In some studies, response to first capture does occur. A "trap happy" animal may be attracted to bait used in trapping and be recaptured more easily after the first capture. A "trap shy" animal may have a reduced probability of capture after being stressed by the first capture (i.e., the animal is not available for capture because it is recuperating or the animal is avoiding the traps). The conditional binomial removal estimator (Moran 1951, Zippin 1956) and the truncated geometric removal estimator (Seber 1982) use data from the first capture (removal data). They do not depend on recaptures and are not affected by behavioral response, but they require assumptions that probabilities of first capture are constant through time and constant among animals.

In this dissertation we address estimation of size of a closed population with heterogeneous capture probabilities. In the notation of Otis et al. (1978), this corresponds to models M_h and M_{bh}. Model M_h has a different capture probability for each animal in the population though some of these animals may not be captured during a study. The resulting model contains too many parameters to estimate. To reduce the number of parameters, Burnham (1972) developed a hierarchical model where capture probabilities of individuals come from a beta distribution, but this model was not effective for estimation of N. Burnham also developed a jackknife method (Burnham 1972, Burnham and Overton 1978) that has been very popular. To allow a different capture probability for each animal, Pollock (Otis et al. 1978) developed a multinomial model, M_{bh}. From M_{bh}, Otis et al. (1978) developed the generalized removal (GR) estimator which is an extension of the Zippin estimator. They transform the parameters of model M_{bh} into a new set of parameters $\overline{p}_1, \overline{p}_2, ..., \overline{p}_k$. Otis et al. (1978) describe $\overline{p}_j$ as "a

conditional probability that represents the average first capture probability of those members of the population that have not yet been captured at the time of the jth trapping occasion." These parameters relate to capture probabilities, but are not explicit representations of the capture probability of an individual animal. Also, for the conditions of model M_{bh}, Pollock and Otto (1983) developed $\widehat{N}_{Jb}$, a jackknife version of the number of animals seen (M). The generalized removal estimator and $\widehat{N}_{Jb}$ do not model the trapping process with explicit interpretation of parameters. Norris and Pollock (1996) used nonparametric maximum likelihood estimation of a binomial mixture distribution for M_h and a censored geometric distribution for M_{bh}.

Young et al. (1952) shows heterogeneity in recapture probabilities of house mice. Current methods of estimating N do not use recapture data for characterizing heterogeneity in the presence of behavioral response to first capture. If one assumes that the grouping of animals in a population according to their common probability of first capture stays the same after first capture, then recapture data can help in estimating the mixture and N. In Chapter 2 for removal data, we develop geometric finite mixture models and provide examples of their use on ecological data. In Chapter 3 for mark-recapture data, we combine a binomial finite mixture for recaptures with a geometric finite mixture for first captures to better estimate mixing proportions. Chapter 4 contains simulations for comparison of our estimators to other estimators.

CHAPTER 2. REMOVAL

2.1 Introduction

One approach to estimation of population size is to model the first captures as if the data come from a removal experiment. Zippin (1956) developed a multinomial model for the number of animals first captured at each trapping occasion. The assumptions for this model are that the population is closed, capture probabilities are constant in time and among animals, and capture events are independent (i.e., the capture of one animal does not affect the capture of another animal). As an alternative, Seber (1982, page 498) gives a truncated geometric probability model for the time of first capture for each animal observed. This model requires assumptions similar to those of the Zippin model, but provides a different probability framework. For one animal on one trapping occasion, the probability of capture is p and the probability of not being caught is $1-p$. For a series of independent trapping occasions, the probability of first being caught on occasion x is $p(1-p)^{x-1}$. Using the Seber framework based on time of first capture, this chapter will develop geometric finite mixtures for estimation of population size (N) when capture probabilities among animals are heterogeneous.

In some mark-recapture studies of animal populations, animals have a behavioral response to their first capture. A "trap happy" animal may be attracted to bait used in trapping and be recaptured more easily after the first capture. A "trap shy" animal may have a reduced probability of capture after being stressed by the first capture (i.e., the animal is not available for capture because it is recuperating or the animal is avoiding the traps). Removal estimators, such as those of Zippin (1956) and Seber (1982), use data from the first capture. They do not depend on recaptures and are not affected by behavioral response.

Capture probabilities may not be constant among animals. Animals may vary in their

movements, attraction to bait and traps, or knowledge of the presence of traps. For example, small mammal biologists often prebait their trapping grid by leaving baited traps propped open for some days before trapping begins. Animals that encounter the bait before trapping starts will have a greater probability of capture than those that have not yet found the bait. To allow a different capture probability for each animal, Pollock (Otis et al. 1978) developed a multinomial model, M_{bh}. Relaxing the assumption of constant capture probability by assigning each animal a parameter for capture probability may not be necessary. Finite mixture models can relax the assumption by allowing each animal to belong to one of a few groups that have different capture probabilities without requiring a large number of parameters.

From M_{bh}, Otis et al. (1978) developed the generalized removal (GR) estimator which is an extension of the Zippin estimator. They transform the parameters of model M_{bh} into a new set of parameters $\overline{p}_1, \overline{p}_2, ..., \overline{p}_k$ (k = the last trapping occasion). Otis et al. (1978) describe $\overline{p}_j$ as "a conditional probability that represents the average first capture probability of those members of the population that have not yet been captured at the time of the jth trapping occasion." These parameters relate to capture probabilities, but are not explicit representations of the capture probabilities of individual animals. Also, under the conditions of model M_{bh}, Pollock and Otto (1983) developed $\widehat{N}_{Jb}$, a jackknife version of the number of animals seen (M). We will develop geometric finite mixture models of the capture process which have parameters with explicit interpretations.

2.2 Model Development

A common theme among the models developed in this chapter is that estimation of population size may be accomplished using time of first capture of individual animals as the fundamental response variable of interest. For a population of size N, let $\mathbf{X} \equiv \{X_i : i = 1, \ldots, N\}$ denote the times of first capture of the animals in the population. This framework assumes that, if trapping were continued indefinitely, the probability that an animal in the population could remain uncaptured would become negligible. A convenient distribution to model the probability distribution of the X_i is the geometric distribution. For $i = 1, \ldots, N$, let X_i have

probability mass function

$$g(x; p_i) = p_i(1-p_i)^{x-1}; \quad x = 1, 2, \ldots \tag{2.1}$$

for an unknown parameter $0 < p_i < 1$; $i = 1, \ldots, N$. Here, p_i represents the probability of capture for a given animal i, and is assumed constant across time for each animal. A basic geometric model would take $p = p_1 = \ldots = p_N$ and assume that $\{X_i : i = 1, \ldots, N\}$ are independent, yielding

$$g(x; p) = p(1-p)^{x-1}; \quad x = 1, 2, \ldots \tag{2.2}$$

as the mass function of each X_i; $i = 1, \ldots, N$.

2.2.1 A Homogeneous Geometric Model

Population surveys do not run until all animals in a population are captured, so that defining X_i to represent observable quantities with support $\{1, 2, \ldots\}$ is not realistic. A study will typically involve a finite set of discrete trapping occasions so that a random variable connected with the time of first capture for animal i may range from 1 to k, where k is the number of trapping occasions. For the M number of animals captured, let $\mathbf{Y} \equiv \{Y_i : i = 1, \ldots, M\}$ denote the times of first capture given that the animals are caught at least once in k samples. Times of first capture greater that k are not observed. This truncation problem is similar to right censoring in survival analysis, but it is different in that the number of animals available for capture is not known. At the end of a survival study, the number of subjects at risk of failure is known. At the end of a removal survey, the number of animals at risk of capture is unknown. The random variables $\mathbf{X}$ and $\mathbf{Y}$ explain the same phenomenon, but differ in their support. The support of Y_i is $\{1, 2, \ldots, k\}$. The random variables $\mathbf{X}$ are on the whole population. Let p be the common capture probability for the N animals in the population. By the assumption that p is the same value for captured and uncaptured animals, inferences can be made about the uncaptured animals. Estimation of N (Section 2.3) is based on $\mathbf{X}$. Observations of $\mathbf{Y}$ are used to estimate p. A distribution that may be used to model time of first capture in this typical setting of k trapping occasions is a truncated version of (2.2),

giving probability mass function (pmf) for Y_i; $i = 1, 2, \ldots, M$,

$$f_{Y,1}(y; p) = \frac{p(1-p)^{y-1}}{1-(1-p)^k}; \quad y = 1, 2, \ldots, k, \tag{2.3}$$

where the denominator in (2.3) is a normalizing constant that adjusts for truncation of the support of Y_i at trapping occasion k. This pmf (2.3) is equivalent to equation (12.54) of Seber (1982).

For a population survey consisting of k trapping occasions and resulting in a record of M times of first capture $\mathbf{y} \equiv (y_1, \ldots, y_M)^T$, the log likelihood formed from mass functions (2.3) is

$$L(p|\mathbf{y}) = \sum_{i=1}^{M} \left[\log(p) + (y_i - 1)\log(1-p) - \log\{1-(1-p)^k\}\right] .$$

This log likelihood is conveniently represented as a function of the number of animals with first-time captures on each of the k trapping occasions. That is, denote the number of first-time captures on trapping occasion j as $m_j \equiv \sum_{i=1}^{M} I(y_i = j)\,; j = 1, \ldots, k$, where $I(\cdot)$ denotes the indicator function having value 1 if the associated argument is true and 0 otherwise. Then the log likelihood for a homogeneous geometric model may be written as,

$$L(p|\mathbf{y}) = \sum_{j=1}^{k} \{j\, m_j \log(1-p)\} + M\log(p) - M\log(1-p) - M\log\{1-(1-p)^k\} . \tag{2.4}$$

2.2.2 A Geometric Finite Mixture Model

The homogeneous geometric model of section 2.2.1 accounts for the finite duration of a population survey, but does not allow heterogeneity in capture probabilities among animals. When some animals have a greater capture probability than others in the population, they tend to be caught in the early occasions of the survey more than animals with smaller capture probabilities. A model that allows this heterogeneity may be formed as a finite mixture of geometric mass functions. To accomplish this, assume that animals in the population may be conceptualized as belonging to one of two groups with differing capture probabilities between the groups. Let the proportion of animals from the entire population that belong to the first group be π, and assume that these animals have capture probability p_1. Similarly, let $(1-\pi)$ denote the proportion of animals in the second group, having capture probability p_2. Then

the probability mass function for a randomly chosen X_i is

$$f_{X,2}(x;\pi,p_1,p_2) = \pi g(x;p_1) + (1-\pi)g(x;p_2)\,; \quad x = 1,\, 2,\, \ldots, \tag{2.5}$$

where $g(\cdot)$ is given in (2.2). This mass function is in the form of a finite mixture consisting of two component distributions (e.g., Everitt and Hand 1981). With truncation of the survey on occasion k, the pmf for a randomly chosen Y_i is

$$f_{Y,2}(y;\pi,p_1,p_2) = \frac{f_{X,2}(y;\pi,p_1,p_2)}{F_{X,2}(k;\pi,p_1,p_2)}\,; \quad y = 1,\, 2,\, \ldots,\, k. \tag{2.6}$$

For a set of M observed times of first capture, the log likelihood formed from mass function (2.6) is simply

$$L(\pi,p_1,p_2|\mathbf{y}) = \sum_{i=1}^{M} \log\{f_2(y;\pi,p_1,p_2)\} - M\log\left\{F_2\left(k;\pi,p_1,p_2\right)\right\}\,,$$

which may again be written in terms of the number of first-time captures at each trapping occasion as

$$\begin{aligned} L(\pi,p_1,p_2|\mathbf{y}) = \; & \textstyle\sum_{j=1}^{k} m_j \log\{\pi g(j;p_1) + (1-\pi)g(j;p_2)\} \\ & -M\log\left\{F_2\left(k;\pi,p_1,p_2\right)\right\}, \end{aligned} \tag{2.7}$$

where $F_2\left(\cdot\right)$ is the cumulative distribution function (cdf) corresponding to (2.5). The last term in (2.7) adjusts the log likelihood for truncation of sampling at occasion k.

2.2.3 A Restricted Geometric Mixture Model

In some capture surveys, the number of captures is much greater on the first trapping occasion than on subsequent occasions. In the context of the finite mixture model of Section 2.2.2, this corresponds to a very high capture probability for one of the two groups of animals. Cohen (1966) describes similar situations for problems involving infection in which the class of non-infected individuals is large. One way to model the occurrence of a large number of first-time captures on the first trapping occasion is to assume that all animals in one of the groups in a finite mixture model are captured. Here, we will take these animals to be in group 2, and set the capture probability for that group to be 1 (i.e., $p_2 = 1$). The finite mixture

model (2.5) then becomes

$$f_3(x;\pi,p_1) = \begin{cases} \pi p_1 + (1-\pi)\,; & x = 1 \\ \pi g(x;p_1)\,; & x = 2, \ldots, \end{cases} \tag{2.8}$$

since, from (2.2), $g(1;p) = p$ for any p, and $g(x;1) = 0$ for $x = 2, \ldots$. The log likelihood for a truncated version of (2.8) is

$$\begin{aligned} L(\pi,p_1|\mathbf{y}) \;=\; & \sum_{i=1}^{M} \log\left[\{\pi p_1 + (1-\pi)\} I(y_i = 1) + \pi g(y_i;p_1) I(y_i \geq 2)\right] \\ & -M \log\left\{F_3\left(k;\pi,p_1\right)\right\}, \end{aligned}$$

or, again using m_j to denote the number of first-time captures on trapping occasion j,

$$\begin{aligned} L(\pi,p_1|\mathbf{y}) = \; & m_1 \log\{\pi p_1 + (1-\pi)\} + \textstyle\sum_{j=2}^{k} m_j \log\{\pi g(y_i;p_1)\} \\ & -M \log\left\{F_3\left(k;\pi,p_1\right)\right\}. \end{aligned} \tag{2.9}$$

In the last term of the log likelihood (2.9), $F_3(\cdot)$ is the cdf of (2.8) and is used to account for truncation at occasion k.

2.3 Estimation

Consider any model of time of first-capture in a closed population of N animals, with N being unknown and the primary focus of estimation. Under the assumption that if trapping were continued indefinitely all animals would eventually be caught, we may consider the elements of $\mathbf{X} \equiv \{X_i : i = 1, \ldots, N\}$ to be independent and identically distributed with support on the set of non-negative integers that represent hypothetical trapping occasions. The observed number of unique animals caught in k trapping occasions, for fixed k, may then be considered a random variable, M_N say. Then $M_N = \sum_{i=1}^{N} I\left(X_i \leq k\right)$ and, $E_X(M_N) = NPr(X_i \leq k) = NF(k;\theta)$, where F is the common distribution function of the random variables $X_i; i = 1, \ldots, N$, and depends on the parameter θ. If θ is estimated by the method of maximum likelihood, a maximum likelihood estimate of the population size is given by

$$\hat{N} = \frac{M}{F(k;\hat{\theta})}, \tag{2.10}$$

where M is the observed value of M_N. In the case of a geometric model (2.2) with common parameter p for time of first capture in the population, this leads to the estimator of Seber (1982),

$$\hat{N} = \frac{M}{1-(1-\hat{p})^k}. \tag{2.11}$$

Trapping is conducted only a finite k number of times, thus making observation of any values $X_i > k$ impossible. For this reason, estimation of θ is based on the random variables $\mathbf{Y}$, which are observable and have the support of $\{1,\ldots,k\}$. The probability mass functions for $\mathbf{Y}$ are versions of those for $\mathbf{X}$ with adjustments for truncation. For example, the truncated homogeneous geometric model (2.3) is simply a censored version of (2.2). The log likelihood functions (2.4), (2.7), and (2.9) for the truncated homogeneous geometric, geometric finite mixture, and restricted geometric mixture models, respectively, lend themselves to maximization by standard iterative procedures such as Newton-Raphson. The EM algorithm (Appendix A) is one alternative method of maximization. The resulting maximum likelihood estimate (MLE) $\hat{\theta}$ is then used in (2.10) to produce an MLE of population size. Care should be taken to verify that estimates are global MLE's. When searching for the MLE's of the geometric finite mixture model, some starting values lead to convergence at a local maximum corresponding to $p_2 = 1$ or $p_1 = p_2$. The standard error of $\hat{N}$ may be computed by the delta method for transformation of asymptotically normal statistics. For the examples of Section 2.4 in this chapter, the covariance matrix of $\widehat{\theta}$ was calculated as the inverse observed information matrix (Appendix B).

2.4 Examples

Data from studies on meadow vole, house mouse, northern pike, and pocket mouse were analyzed to illustrate the use of the geometric mixture models. These models do not require that all animals have the same capture probability, however, they require assumptions that the parameters are constant in time (i.e., population closure and constant capture probabilities). For example, animal captures should not be affected by full traps or changes in the weather.

Model selection is accomplished using Pearson GOF tests and likelihood ratio (LR) tests.

The homogeneous geometric model is nested within the restricted geometric mixture model by using the constraint $\pi = 1$. The restricted geometric model is nested within the geometric finite mixture model by using the constraint $p_2 = 1$. The geometric finite mixture model is referred to as the full model, because all parameters are included in the model without constraint. Generalized removal estimates and the Pollock and Otto jackknife estimate are calculated with the program CAPTURE (Rexstad and Burnham 1991).

2.4.1 Meadow Vole

Small mammal biologists often use a prebait period before actual trapping starts, leaving traps propped open with bait for 2 days. This acts to increase the number of animals caught in the first few days of trapping, but it may introduce heterogeneity in the capture probabilities among animals in that animals familiar with the bait have a higher capture probability than those that have not yet found the bait. To illustrate this problem, we analyzed data from a mark-recapture survey on meadow vole (*Microtus pennsylvanicus*) at Doolittle Prairie, Story County Iowa. A 10 by 10 grid with 12 meter spacing of Sherman live traps was baited with oats and traps were left open for 2 days prior to 5 days (occasions) of sampling during 21-25 June 1994. The traps were checked 3 times during each trapping occasion to ensure the availability of traps. When a trap site was checked, if no open trap was available at the site, then another baited open trap was placed at the site to ensure capture opportunity for voles in the vicinity. The frequencies of first time captures were 48, 12, 8, 9,and 6 for days 1 through 5, respectively.

Table 2.1 lists estimates for the geometric models of which the full finite mixture and the restricted mixture have good fits (GOF ≤ 0.46, P > 0.4). The likelihood ratio test comparing these models is not significant (P $= 0.7$). Therefore, the more parsimonious, restricted geometric mixture model is used for estimation ($\widehat{N} = 112$, se $= 25$). Figure 2-1 presents data and expected frequencies for geometric models. Figures 2-1a and 2-1b look alike, however estimates of N differ by 13% (Table 2.1). Differences in the models are reflective in the way that expected capture frequencies approach 0 as occasions increase. Expected values of the restricted mixture model (Table 2.2, Figure 2.1A) approach 0 faster than those of the finite

Table 2.1

*Estimates for geometric models on meadow vole first capture data, k=5 (*GOF = *Pearson goodness of fit chisquare,* L = *log likelihood).*

	Model		
	Full	Restricted	Homogeneous
N (se)	129 (110)	112 (33)	87 (2.2)
95% CI	(83, 351)	(83, 178)	(83, 92)
GOF	0.39	0.46	10.03
L	-104.157	-104.199	-109.218
π	0.71	0.697	1
p_1	0.127	0.178	0.452
p_2	0.948	1	-

Table 2.2

Expected first capture frequencies for full and restricted geometric mixture models on meadow vole data, k=5.

		expected	
occasion	observed	full	restricted
1	48	48.0	48.0
2	12	11.9	11.5
3	8	8.8	9.4
4	9	7.6	7.7
5	6	6.7	6.4
6		5.8	5.2
7		5.1	4.3
8		4.4	3.5
9		3.9	2.9
10		3.4	2.4
15		1.7	0.9
20		0.9	0.3
25		0.4	0.1

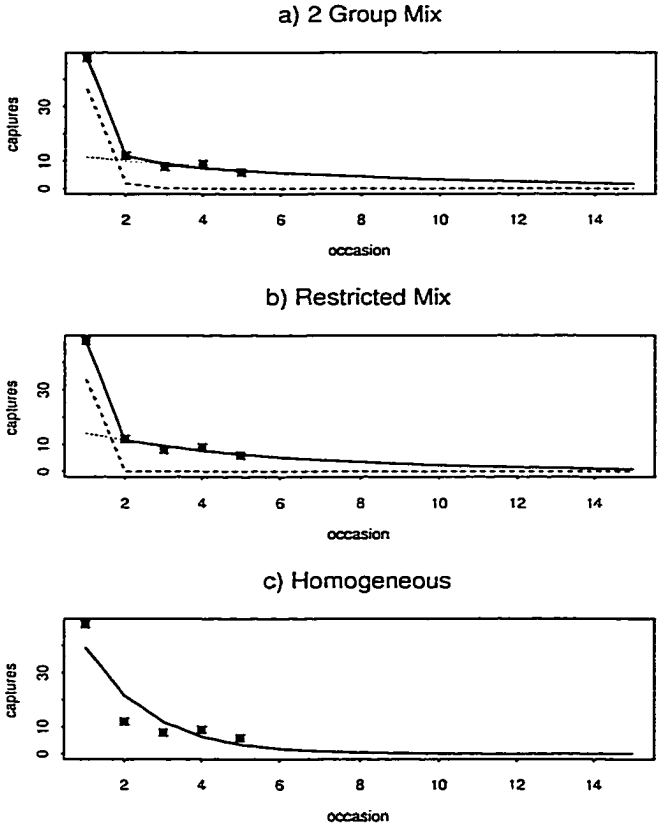


Figure 2.1 Observed (asterisks) and expected frequencies for geometric models fit to meadow vole data. Solid, dot, and dash lines represent the mixture, group 1, and group 2, respectively.

mixture model (Table 2.2, Figure 2.1B) as occasions increase. The area under the curve after occasion 5 relates to the estimate of uncaptured voles. This area is greater for the full mixture model.

For the generalized removal (GR) model, the population size estimate is $\widehat{N} = 102$ (se = 20) with $\widehat{\overline{p}}_1 = 0.471$ and $\widehat{\overline{p}}_2 = \widehat{\overline{p}}_3 = \widehat{\overline{p}}_4 = \widehat{\overline{p}}_5 = 0.226$. The Pollock and Otto jackknife estimate, $\widehat{N}_{Jb}$ is 107 (se = 11). From simulation results on the GR estimator (Otis et al. 1978) and expected values of $\widehat{N}_{Jb}$, Pollock and Otto (1983) found that $\widehat{N}_{Jb}$ has less negative bias than the GR estimator under high heterogeneity. The GR estimate and $\widehat{N}_{Jb}$ are similar to the restricted mixture estimate, however the GR model suggests that uncaptured voles have a capture probability of 0.226 which is significantly larger than 0.178 estimated with the restricted mixture model.

2.4.2 House Mouse Recapture

Young et al. (1952) studied the heterogeneity of trap response in 2 populations of house mouse (*Mus musculus L.*) living as natural infestations in heated buildings. Live-traps were operated for 3 consecutive nights each week followed by a four-day rest period. The study period extended from September 28, 1948 to May 27, 1949. On its original capture, each mouse was marked with an individual pattern of toe-clipping and released. They estimated a recapture probability of 0.46 for a homogeneous geometric distribution from the length of time elapsing between the initial capture and the first recapture of 721 mice. They found that this model had a poor fit (GOF $\chi^2 = 14$, P< 0.0001). With this information and the pattern of deviance, they concluded that recapture probabilities were heterogeneous. The approximate frequencies for times elapsing between the initial capture and the first recapture are 433, 113, 57, 37, 28, 7, 13, 13, 2, 5, 2, 0, 2, and 2 for lapses 1 through 14 occasions, respectively (from Figure 2.4 of Young et al., 1952 page 177). Though these data are on a geometric process of recaptures, the data are compatible with the geometric finite mixture models developed here. We use these models to characterize heterogeneity in recapture probabilities.

Of the three models, the geometric finite mixture has a reasonable fit (GOF= 16.35, P=

Table 2.3
Estimates for geometric models on house mouse recapture data, k=14 (GOF = *Pearson goodness of fit chisquare,* L = *log likelihood*)

	Model		
	Full	Restricted	Homogeneous
N (se)	716 (0.87)	715 (0.36)	714 (0.029)
95% CI	(714, 718)	(714, 716)	(714, 714)
GOF	16.35	23.79	190.23
L	-987.16	-989.67	-1037.26
π	0.482	0.609	1
p_1	0.311	0.352	0.476
p_2	0.878	1	-

0.09, Table 2.3). The plot of the homogeneous geometric model (Figure 2.2c) is similar to that of Young et al. (1952). The likelihood ratio tests for comparing the full geometric finite mixture to the restricted mixture and to the homogeneous geometric model were significant (P= 0.025 and P < 0.001, respectively). This indicates that the geometric finite mixture is the best model. This choice is consistent with the heterogeneity in recapture probabilities noted by Young et al. (1952).

Groups 1 and 2 have estimated recapture probabilities of 0.311 and 0.878, respectively (Table 2.3). Mice in group 2 are expected to be "removed" sooner than group 1, because of the higher recapture probability. The estimated probabilities of being in group 1 given y are 0.248, 0.650, 0.913, 0.983, and 0.997 for y equal to 1 through 5, respectively. Heterogeneity among uncaptured animals is greatest on the first occasion, then gradually declines. Figure 2.2a shows that almost all mice with a lapse greater than 3 occasions are modeled as belonging to group 1. After 3 occasions, expected values for group 1 are nearly the same as the mixture. At this point, the mice are nearly homogeneous with the same recapture probability.

Most mark-recapture data sets are shorter than the mouse recapture data above. To demonstrate estimation with less data, a second data set with time elapsed between initial capture and first recapture was formed by truncating the mouse data to lapses of 1 through 5 occasions. All three models were fit to these data (Table 2.4, Figure 2.3). The restricted geometric mixture had the best fit of the 3 models. The homogeneous geometric model had a

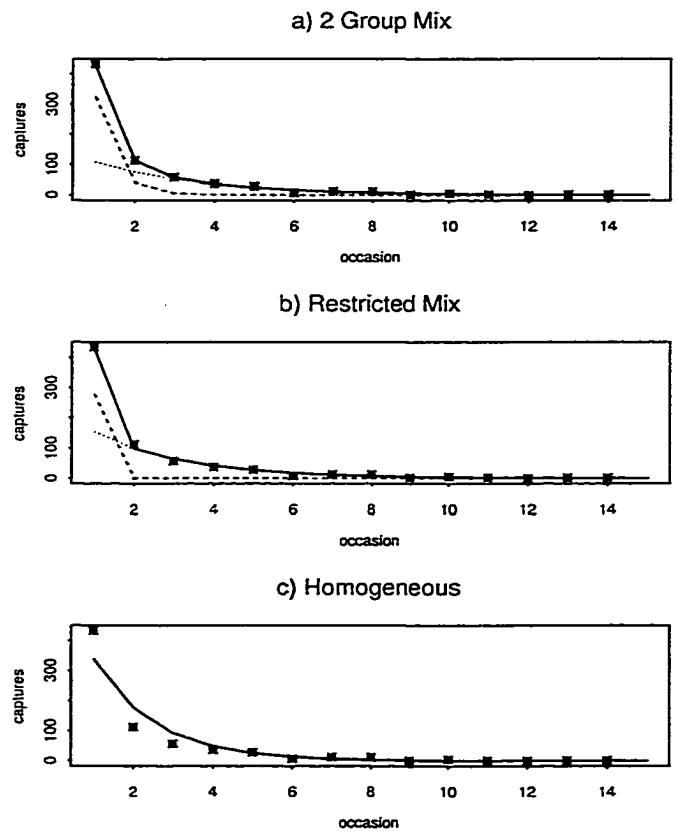


Figure 2.2 Observed (asterisks) and expected frequencies for geometric models fit to house mouse data (k=14). Solid, dot, and dash lines represent the mixture, group 1, and group 2,respectively.

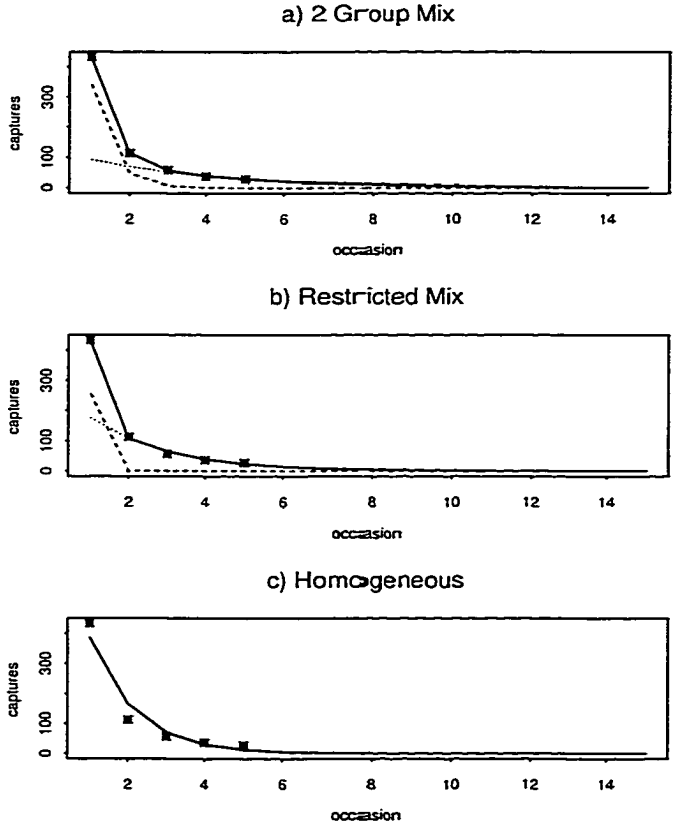


Figure 2.3 Observed (asterisks) and expected frequencies for geometric models fit to truncated house mouse data(k=5). Solid, dot, and dash lines represent the mixture, group 1, and group 2,respectively.

Table 2.4
Estimates for geometric models on house mouse recapture data k=5 (GOF = *Pearson goodness of fit chisquare,* L = *log likelihood*)

	Model		
	Full	Restricted	Homogeneous
N (se)	747 (68)	705 (11)	678 (2.0)
95% CI	(668, 879)	(683, 726)	(674, 682)
GOF	0.04	2.15	43.78
L	-724.70	-725.76	-745.38
π	0.472	0.636	1
p_1	0.260	0.394	0.573
p_2	0.867	1	-

poor fit (GOF P< 0.001). The likelihood ratio test comparing the geometric finite mixture and the restricted geometric mixture models was not significant (P = 0.146). Of these 2 models, the restricted geometric mixture is more parsimonious. From this model, the estimate of the number of marked mice is 705 with a 95% CI of (684, 725). Young et al. (1952) report that 721 mice were recaptured at least once. The generalized removal estimate is 703 (684, 735). The Pollock and Otto jackknife estimate is 780 with a 95% confidence interval (742, 836) that does not include 721.

The long series data set ($k = 14$), the LR and GOF tests indicate that the full model is the best choice, but for the short data set ($k = 5$), the restricted model is the best choice. The choice of model for a data set is in part a function of the capture process that generates the data (i.e., the true value of capture probabilities and the degree of heterogeneity). It is also a function of how much information is available. This example shows that a long data series has more information and can allow greater precision and discrimination between models.

2.4.3 Northern Pike

Gillnets and seine hauls were used to capture northern pike at Deadman lake, a 213 ha lake 100 km southwest of Fairbanks Alaska (Hansen and Pearse 1995). The daily sampling protocol did not change for 5 consecutive days of sampling. Daily frequencies were 261, 149, 119, 73, and 39 for first time captures on days 1 to 5, respectively. All three models fit well (Table 2.5,

Table 2.5
Estimates for geometric models on northern pike first capture data, k = 5 (GOF = Pearson goodness of fit chisquare, L = log likelihood).

	Model		
	Full	Restricted	Homogeneous
N (se)	730 (21)	730 (22)	722 (14)
95% CI	(690, 771)	(687, 773)	(694, 750)
GOF	3.68	3.68	4.04
L	-921.9	-921.9	-922.1
π	0.973	0.973	1
p_1	0.340	0.339	0.355
p_2	1	1	-

Figure 2.4), but the homogeneous geometric model is most parsimonious (LR restricted mix vs homogeneous = 0.4, P= 0.5). The population estimate is 722 (se = 14). Figures 2-4a and 2-4b show that the resulting mixtures are very similar to the homogeneous model in Figure 2.4c, though each model uses the components differently. The generalized removal estimate is 720 (se = 17) and the Pollock and Otto jackknife estimate is 797(se = 28).

2.4.4 Pocket Mouse

Larsen (as reported in Otis et al. 1978) conducted a capture-recapture study on pocket mouse for 7 consecutive nights. The frequencies of first captures were 23, 9, 3, 5, 6, 6, and 3 for occasions 1 through 7, respectively. For these data, all 3 of the geometric models have GOF tests that indicate an adequate fit (Table 2.6, P > 0.08). The LR test between the restricted and homogeneous models is significant (LR = 6.94, P = 0.008) and the LR test between the full and restricted models is not significant (LR = 0.968, P = 0.3), therefore the restricted model is the best of choice of the 3 ($\widehat{N} = 90$). Figure 2.5 shows the fit of these models.

These data have a bimodal pattern with the second mode occurring at occasions 5 and 6. This pattern could arise by 1) chance under a geometric removal process (expected frequencies monotonically decrease), 2) capture probabilities changing in time (e.g., a storm suppressing movement), or 3) immigration. We simulated the geometric removal process to estimate the

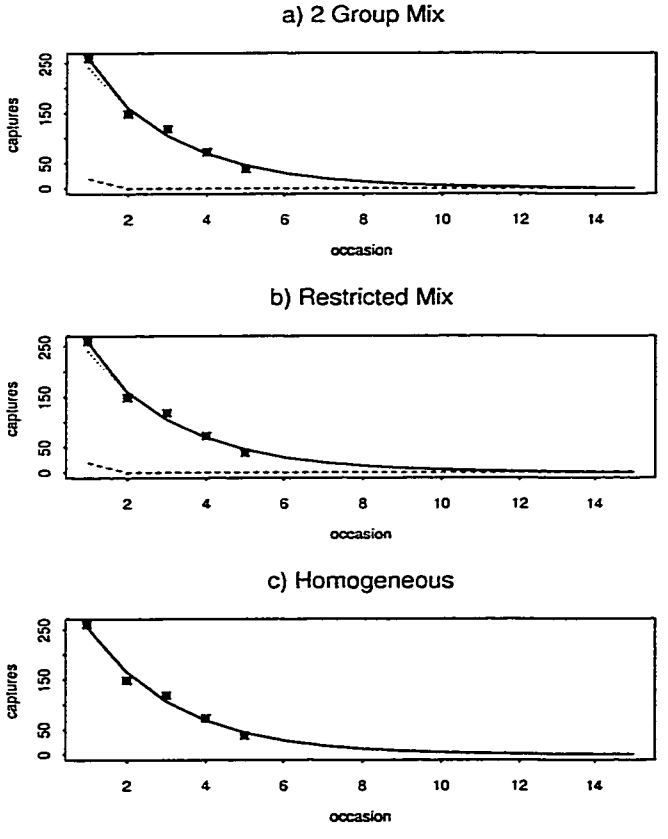


Figure 2.4 Observed (asterisks) and expected frequencies for geometric models fit to northern pike data (k=5). Solid, dot, and dash lines represent the mixture, group 1, and group 2,respectively.

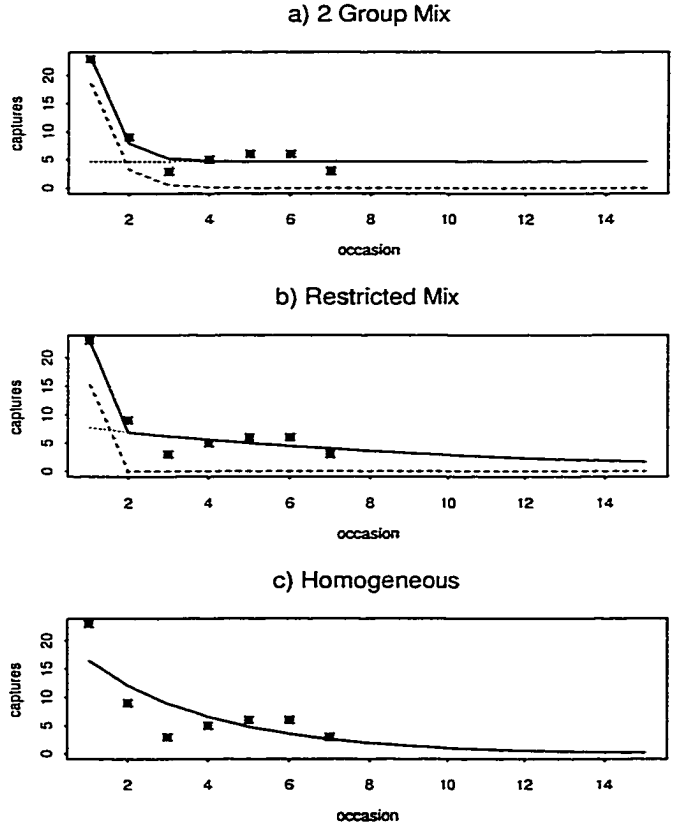


Figure 2.5 Observed (asterisks) and expected frequencies for geometric models fit to pocket mouse data (k=7). Solid, dot, and dash lines represent the mixture, group 1, and group 2,respectively.

Table 2.6

Estimates for geometric models on pocket mouse first capture data, $k = 7$ (GOF = Pearson goodness of fit chisquare, L = log likelihood).

	Model		
	Full	Restricted	Homogeneous
N (se)	170000 (-)	90 (46)	62 (4.3)
95% CI	-	(55, 181)	(55, 71)
GOF	2.48	3.34	9.73
L	-93.698	-94.182	-97.672
π	0.9999	0.830	1
p_1	0.0000266	0.102	0.265
p_2	0.824	1	-

Table 2.7

Simulation results for second mode under a homogeneous geometric model.

p	0.1	0.15	0.2	0.25	0.3
P(second mode $\geq 5 \mid p$)	0.470	0.290	0.157	0.062	0.049

probability of bimodality. We simulated 1000 replicates for the capture of 100 animals under a homogeneous geometric model with no truncation (i.e., $k = \infty$). For each replicate we calculated the mode for occasions 3 and greater. The probabilities of a second mode at occasion 5 or greater are listed in Table 2.7. With $p = 0.1$, this bimodal pattern in the data could have occurred by chance (P = 0.470). The geometric capture process is without memory, future captures are independent of past captures. After occasion 2, the remaining uncaptured mice are expected to be captured with frequencies that follow a geometric decline. However, for occasions 3 through 7, frequencies do not decline until the last occasion. This pattern may contribute to the difficulty in estimation for the full model (Table 2.6, $\widehat{N} = 170,000$).

Norris and Pollock (1996) used nonparametric maximum likelihood estimation (Lindsay and Roeder 1992) under model M_{bh} to estimate a distribution for capture probabilities as having 0.092 of the animal population belonging to a group with capture probability of 0.798 and 0.881 of the population belonging to another group with capture probability 0.202. From their model, $\widehat{N} = 92$ (conditional bootstrap se = 22). Though their model is different, these values correspond to $\pi = 0.798$, $p_1 = 0.092$, and $p_2 = 0.881$ for the geometric finite mixture

model. The log likelihood of the geometric finite mixture evaluated at these values is -93.896, which is not the maximum. In bootstraps and simulations, Norris and Pollock (1996) only used their MLE when $\widehat{N}$ is less than a preset upper bound for N and they acknowledge that the rule is ad hoc. This condition could cause a negative bias in their estimate for the se of $\widehat{N}$.

2.5 Discussion

We have developed geometric finite mixture models for estimation of population size from closed populations. We have modeled time to first capture as if each animal is associated with a discrete group in which members have a common first-capture probability. Model parameters have explicit interpretations relating to the capture process. Our models are parametric and we use maximum likelihood to estimate capture probabilities and mixing proportions. Model selection can be accomplished using standard goodness of fit and likelihood ratio tests. Approximate variances can be calculated by use of the inverse information matrix and the delta method. We have demonstrated the use of the two component geometric finite mixture (full), a restricted mixture ($p_2 = 1$), and the homogeneous geometric model ($p_1 = p_2 = p$) on four removal data sets. The geometric models developed here provide a probability framework for estimating N that is an alternative to multinomial models. If recapture probabilities among animals are correlated with first-capture probabilities, then the estimation of mixing proportions may be enhanced by modeling recapture data with finite mixtures.

For estimation of N, we need to make inferences about animals that did not get caught. According to our models, heterogeneity is greatest on the early occasions, so that animals with large capture probabilities are removed earlier than those with smaller capture probabilities. The heterogeneity among uncaptured animals declines gradually.

Extending a study (i.e., larger k) would provide more information for estimation (see e.g., Section 2.4.2). However, a longer study may risk violating assumptions of closure or constant capture probabilities. Migration, death, or recruitment can compromise the closure of the population. Changes in weather, adverse natural events (e.g., flooding, predation, fire), and natural history events (e.g., breeding) can cause capture probabilities to change through time.

Generally, we want to increase capture probabilities to get greater precision in estimates. In the meadow vole example (Section 2.4.1), traps were baited and open for 2 days prior to trapping in an effort to increase capture probabilities. The result was a mixture of high and low capture probabilities. We needed more parameters to model these data than if probabilities had been homogeneous. Thus, precision of estimates may have been reduced through the introduction of heterogeneity caused by the trapping protocol.

Heterogeneity in capture probabilities among animals in a population may be intrinsic to the animals. Before a survey starts, animals may differ in tendency to move or in desire for the bait. Heterogeneity may also be induced by features of the sample design. Perhaps heterogeneity is a function of an animals home range and the position of a trap site. An animal close to a trap would have an opportunity for capture without much search time. An animal farther away from a trap may not encounter a trap until it moves in search of food or shelter. This additional travel time may contribute to lower capture probabilities for some animals. In this case, increasing trap density may serve to reduce heterogeneity and increase capture probabilities.

Ideally, a capture survey would be free of heterogeneity as is apparently the case for the northern pike example (Section 2.4.3). Gillnets can readily catch multiple fish. A fish in the net does not deter other fish from being captured. In contrast, for small mammal surveys traps typically hold only one animal. Recaptured animals may occupy traps that would otherwise provide opportunity of first-capture. This interference may be a source of heterogeneity and is more likely when trap density is low or when animals are highly aggregated.

A common sampling method is to spread traps uniformly across the sample area, though animals are typically not uniformly spread. If concentrations of animals have capture opportunities limited by nearness of open traps, then capture events are unlikely to be independent. Adding traps to positions that have caught animals may serve to ensure availability of traps in an adaptive sampling fashion, but may create a change in effort that requires a new model. The restricted geometric mixture model adjusts for inflated frequencies of captures on occasion 1. Perhaps, a model that likewise controls for suppression of captures on occasion 1 would model the increased effort after occasion 1.

CHAPTER 3. MARK-RECAPTURE

3.1 Introduction

In Chapter 2, animals in a population were modeled as belonging to one of 2 groups. Between groups, the probability of first capture could be different, but within each group, animals had the same capture probability. We now extend this grouping to model recapture probabilities. Such a model allows heterogeneity in probabilities of first capture and heterogeneity in behavioral response to first capture. Animals within each group have the same probabilities of first capture and the same probabilities of recapture, but between groups, these probabilities are allowed to differ.

In Chapter 2, population size estimates were based on finite mixture models of time to first capture. Recapture events were not considered. This chapter further develops those models to include recapture events in the estimation of population size. The jackknife estimator of Burnham and Overton (1979) uses recapture data, but assumes that for each animal the recapture probability is equal to the probability of first capture. Models in the following section allow each group in the finite mixture to have a recapture probability that differs from the probability of first capture and make use of recapture data in estimation of population size, N.

3.2 Models

For a population of size N, let $\mathbf{X} \equiv \{X_i : i = 1, \ldots, N\}$ denote the times of first capture for all animals in the population. This framework assumes that, if trapping were continued indefinitely, the probability that an animal in the population could remain uncaptured would become negligible. A convenient distribution to model the probability distributions of the X_i

is the geometric distribution. For $i = 1, \ldots, N$, let X_i have probability mass function

$$g(x; p_i) = p_i(1 - p_i)^{x-1}; \quad x = 1, 2, \ldots$$

for an unknown parameter $0 < p_i < 1$; $i = 1, \ldots, N$. Here, p_i represents the probability of capture for a given animal i, and is assumed constant across time for each animal. A basic geometric model would take $p = p_1 = \ldots = p_N$, yielding

$$g(x; p) = p(1 - p)^{x-1}; \quad x = 1, 2, \ldots \tag{3.1}$$

as the mass function of each X_i; $i = 1, \ldots, N$.

3.2.1 Homogeneous Model

Population surveys do not run until all animals in a population are captured, so that defining X_i to represent observable quantities is not realistic. A study will typically involve a finite set of discrete trapping occasions so that a random variable connected with the time of first capture for animal i may range from 1 to k, where k is the number of trapping occasions. For the M animals captured, let $\mathbf{Y} \equiv \{Y_i : i = 1, \ldots, M\}$ denote the times of first capture, given that the animals are caught at least once in k samples. Times of first capture greater than k are not observed. This truncation problem is similar to right censoring in survival analysis, but it is different in that the number of animals available for capture is not known. At the end of a survival study, the number of subjects that have not failed is known. At the end of a mark-recapture survey of a population, the number of animals that have not been captured, $N - M$, is unknown and is, in fact, the focus of estimation. The random variables $\mathbf{X}$ may be used to model the entire population, while the random variables $\mathbf{Y}$ may be used to model that portion of the population that is actually captured. Under the assumption that p, the common capture probability for the N animals in the population, is the same for captured and uncaptured animals, estimation of N is based on the model for $\mathbf{X}$, while observations of $\mathbf{Y}$ from the first k occasions are used to estimate p.

Given that animal i was first captured on occasion j, its recapture history is a sequence of $d_i - j$ independent Bernoulli outcomes, where d_i is the last occasion available for recapture.

Typically, $d_i = k$, but death of animal i causes right-censoring where $d_i < k$. The random variable for the number of recaptures of animal i, W_i, can be modeled with a binomial distribution with mass function

$$b\left(w|y;c,d\right) = \binom{d-y}{w} c^w \left(1-c\right)^{d-y-w} \tag{3.2}$$

for $w = 0, \ldots, d-y$, where c is the recapture probability and is assumed to be equal among animals. It is assumed that $\{W_i; i = 1, \ldots, N\}$ are independent random variables. We only observe M recapture counts, but potentially N recapture counts could occur. A distribution that may be used to model time of first capture in this typical setting of k trapping occasions is a truncated version of (3.1). Using (3.1) truncated to have support $\{1, \ldots, k\}$ and (3.2), the joint probability mass function (pmf) for Y_i and W_i; $i = 1, 2, \ldots, M$, is

$$f_{Y,W,1}\left(y,w;p,c,d\right) = \frac{g\left(y;p\right)}{F_g\left(k;p\right)} b\left(w|y;c,d\right) \; ; \quad y = 1, 2, \ldots, k; w = 0, \ldots, d-y. \tag{3.3}$$

The denominator in (3.3) on a homogeneous population (one group) is the geometric cumulative distribution function and is a normalizing constant that adjusts for truncation of the support of $g\left(y;p\right)$ at trapping occasion k.

A population survey consisting of k trapping occasions results in a record of M times of first capture $\mathbf{y} \equiv (y_1, \ldots, y_M)^T$ and M recapture counts $\mathbf{w} \equiv (w_1, \ldots, w_M)^T$. The log likelihood formed from mass function (3.3) is

$$L(p,c|\mathbf{y},\mathbf{w},\mathbf{d}) = -M \log F_g\left(k;p\right) + \sum_{i=1}^{M} \left\{\log g\left(y_i;p\right) + \log b\left(w_i|y;c,d_i\right)\right\}. \tag{3.4}$$

This model accounts for recaptures, but the recapture data do not provide better estimation of p than the homogeneous removal model (Chapter 2). The first partial derivative of the one group log likelihood (3.4) with respect to p does not involve recapture data. This means that $\widehat{N}$ is the same as for the removal model (2.4) using only $\mathbf{y}$. The restriction, $p = c$, can be applied if the assumption of no behavioral response is reasonable. With this restriction, recaptures relate directly to estimation of p and N.

3.2.2 Finite Mixture Models

The homogeneous geometric model of section 2.1 accounts for the finite duration of a population survey, but does not allow heterogeneity in capture probabilities among animals. When some animals have a greater capture probability than others in the population, they tend to be caught in the early occasions of the survey more than animals with smaller capture probabilities. A model that allows this heterogeneity may be formed as a finite mixture of geometric mass functions. To accomplish this, assume that animals in the population may be conceptualized as belonging to one of two groups which differ in probabilities of first capture and in probabilities of recapture. Let the proportion of animals from the entire population that belong to the first group be π, and assume that these animals are alike in their first capture probability p_1 and recapture probability c_1. Similarly, let $(1-\pi)$ denote the proportion of animals in the second group, having capture probability p_2 and recapture probability c_2. Then the 2-group probability mass function for a randomly chosen X_i is

$$f_{X,2}(x;\pi,p_1,p_2) = \pi g(x;p_1) + (1-\pi)g(x;p_2)\,; \quad x = 1, 2, \ldots, \tag{3.5}$$

where $g(\cdot)$ is given in (3.1). This mass function is in the form of a finite mixture consisting of two component distributions (e.g., Everitt and Hand 1981).

Define $\tau(j)$ as the conditional probability of an animal belonging to group 1, given first capture on occasion j,

$$\tau(j) = \frac{\pi p_1 (1-p_1)^{j-1}}{\pi p_1 (1-p_1)^{j-1} + (1-\pi)\, p_2 (1-p_2)^{j-1}}. \tag{3.6}$$

The function $\tau(\cdot)$ is used in EM algorithm estimation of finite mixtures (e.g. Chapter 2 Appendix A.1, Everitt and Hand 1981). We will use $\tau(\cdot)$ as the probability that the number of recaptures for a given animal belongs to the distribution of recaptures for group 1 animals in a binomial finite mixture distribution for recapture counts. Refer to animals first captured on occasion j as 'cohort' j, among which $\tau(j)$ is the proportion expected to belong to group 1. Note that if $p_1 < p_2$, then $\tau(j)$ increases as j increases (i.e., the proportion of animals belonging to group 1 is greater in cohort $j+1$ than in cohort j). By using $\tau(\cdot)$ as a mixing

proportion in modeling recapture counts, recapture data may be used in estimation of π, p_1, p_2, and N.

The random variable for the number of recaptures of an animal first captured on occasion y_i, denoted W_i can be modeled with a 2 component binomial finite mixture distribution having probability mass function

$$b_2(w|y;\theta,d) = \tau(y)\, b(w|y;c_1,d) + [1-\tau(y)]\, b(w|y;c_2,d), \tag{3.7}$$

where $b(\cdot)$ is given in (3.2). In (3.7), $\theta \equiv (\pi, p_1, p_2, c_1, c_2)^T$, and c_1 and c_2 are recapture probabilities for groups 1 and 2, respectively. For all animals in the same cohort, the W's are assumed to be independent and identically distributed. For $p_1 \neq p_2$, each cohort has a different mixture of group 1 and 2 animals available for recapture. Using mass functions (3.5) and (3.7), the joint pmf of Y_i and W_i is

$$f_{Y,W,2}(y,w;\theta,d) = \frac{f_{X,2}(y;\pi,p_1,p_2)}{F_{X,2}(k;\pi,p_1,p_2)} b_2(w|y;\theta,d)$$

and the log likelihood for all captures and recaptures becomes

$$L(\theta;\mathbf{y},\mathbf{w},\mathbf{d}) = -M\log F_{X,2}(k;\theta) + \sum_{i=1}^{M}\{\log f_{X,2}(y_i;\theta) + \log b_2(w_i|y_i;\theta,d_i)\}. \tag{3.8}$$

In (3.8), M is the total number of animals captured in k occasions and $\mathbf{y},\mathbf{w},\mathbf{d}$ are vectors of length M, consisting of observed times of first capture (y), number of recaptures (w), and number of occasions available for recapture (d), for individual animals captured during the study. The cdf of (3.5), $F_{X,2}(k;\theta)$ is a normalizing constant for truncation of $f_{X,2}(\cdot)$ at occasion k.

The two group model (3.8) can be generalized to a v group version with mixing proportions $\pi_1,\ldots,\pi_v$, which are constrained to have values between 0 and 1 and $\pi_v = 1-\sum_{i=1}^{v-1}\pi_i$. Under such generalization, the probability mass function for X_i, $i=1,2,\ldots,N$, is

$$f_{X,v}(x;\theta_v) = \sum_{l=1}^{v}\pi_l p_l (1-p_l)^{x-1}; \quad x=1,2,\ldots \tag{3.9}$$

The pmf for W_i, $i=1,2,\ldots,N$, is

$$b_v(w|x;\theta_v,d) = \sum_{l=1}^{v}\tau_l(x)\, b(w|x;c_l,d), \tag{3.10}$$

with mixing proportions

$$\tau_u(x) = \frac{\pi_u p_u (1-p_u)^{x-1}}{\sum_{l=1}^{v} \pi_l p_l (1-p_l)^{x-1}}; \quad u = 1, \ldots, v. \tag{3.11}$$

Using (3.9) and (3.10), the joint pmf of Y_i and W_i, $i = 1, \ldots, M$, is

$$f_{Y,W,v}(y, w; \theta, d) = \frac{f_{X,v}(y; \theta)}{F_{X,v}(k; \theta)} b_v(w|y; \theta, d).$$

Then the log likelihood corresponding to the v-group model is

$$L(\theta_v; \mathbf{y}, \mathbf{w}, \mathbf{d}) = -M \log F_{X,v}(k; \theta_v) + \sum_{i=1}^{M} \{\log f_{X,v}(y_i; \theta_v) + \log b_v(w_i|y_i; \theta_v, d_i)\}, \tag{3.12}$$

where $\theta_v \equiv (\pi_1, \ldots, \pi_{v-1}, p_1, \ldots, p_v, c_1, \ldots, c_v)^T$ and $F_{X,v}(\cdot)$ is the cdf of $f_{X,v}(\cdot)$. Like the 2-group model, the general v-group model allows the recapture probability of an animal to differ from its probability of first capture. If the animals do not have a behavioral response to trapping, then the parameter space might be constrained so that $p_l = c_l$ for each group $l = 1, \ldots, v$. The estimate of population size is

$$\widehat{N} = \frac{M}{F_{X,v}\left(k; \widehat{\theta}_v\right)}, \tag{3.13}$$

which is similar to the estimator (2.10) of Chapter 2.3. The estimator $\widehat{N}$ is a function of maximum likelihood estimates (MLE's) for θ_v, therefore it is an MLE and is consistent for N (Bain and Englehardt 1989).

Define M_d as the number of marked animals with d as the common last possible occasion for recapture (e.g., due to death, end of the survey). Among the M_d marked animals with common d, the expected number of animals having values of y and w is

$$E_{YW|M_d}[M_{ywd}] = M_d \cdot f_{Y,W,v}(y, w; \theta). \tag{3.14}$$

From (3.14), the estimated expected frequencies of animals with values of y, w, and d were calculated as

$$\widehat{M}_{ywd} = M_d \frac{f_v\left(y; \widehat{\theta}_v\right)}{F_v\left(k; \widehat{\theta}_v\right)} b_v\left(w|y; \widehat{\theta}_v, d\right).$$

These estimated expected values $\widehat{M}_{ywd}$ are used in goodness of fit tests. For each value of d, the number of yw combinations having frequencies is

$$c(d) = \left(\sum_{i=1}^{d} i\right) - 1.$$

As an example, say that a study has values of $d = 5$ and 3. The number of degrees of freedom is $c(5) + c(3) -$ the number of estimated parameters in θ.

3.3 Examples

The first example is from data that we collected on a population of meadow voles, *Microtus pennsylvanicus,* in 1994 at Doolittle Prairie, Story County Iowa. The area is a typical prairie remnant of native tall grasses with some temporary wetlands. We live-trapped this population of voles during one week in June. The second example is drawn from data presented by Carothers (1973) as sampling scheme a, which has been used as a test data set by other researchers (Otis et al. 1978).

Maximum likelihood estimation of model parameters was accomplished using the direct search algorithm by Hooke and Jeeves (1961). Models were compared using Pearson goodness of fit (GOF) tests, likelihood ratio (LR) tests, Akaike's Information Criteria (AIC), and plots showing fit of models to data. Heterogeneity in capture probabilities among animals is determined by the fit of models. If capture probabilities are reasonably homogeneous, then a 1-group model will be most parsimonious among models with good fit. The asymptotic covariance matrix for $\widehat{\theta}_v$ was calculated as the inverse observed information matrix (Appendix A). The variance of $\widehat{N}$ was calculated using the delta method.

3.3.1 Meadow Vole

During 21-25 June 1994 we trapped meadow voles using Sherman live-traps baited with oats. Live-traps were arranged in a 10 x 10 grid with 12m spacing (1.1664 ha square grid). Traps were baited and opened for 2 days prior to sampling which then continued for 5 days. Once capturing began, traps were set at about 8 pm and checked 3 times during each night

Table 1.

Vole data, n_{ywd} = frequency of animals with corresponding values of d, y, and w.

$d=5$			$d=3$			$d=1$		
y	w	n_{ywd}	y	w	n_{ywd}	y	w	n_{ywd}
1	0	10	1	0	0	1	0	3
1	1	12	1	1	2			
1	2	11	1	2	2			
1	3	9	2	0	0			
1	4	6	2	1	0			
2	0	1	3	0	1			
2	1	4						
2	2	5						
2	3	2						
3	0	3						
3	1	3						
3	2	2						
4	0	5						
4	1	4						
5	0	6						

(occasion) to maintain availability of open traps to uncaptured animals in the vicinity. At each check, an additional baited trap was placed at positions where traps were closed. Except for the last check of the night, all positions had open traps after each check. Captured adult voles were tagged in both ears to avoid problems with tag loss. The capture data are summarized in the Table 1. The number of marked animals is 83 with 3 deaths on occasion 1 due to heat and 5 deaths on occasion 3 due to flooding of trap sites.

We fit 2-group and 3-group full models allowing a behavioral response (no restriction on the parameter space) and 2-group and 3-group restricted models ($\mathbf{p} = \mathbf{c}$) with no behavioral response (Table 2). The LR test comparing the 3-group full and restricted ($\mathbf{p} = \mathbf{c}$) models suggests that the full model is better (LR =10.0294, df = 3, P = 0.018). The LR test comparing the 3-group and 2-group full models suggests that the 3-group full model is better (LR = 9.8294, df = 3, P = 0.02). The LR test comparing the 3-group full model to the 2-group restricted ($\mathbf{p} = \mathbf{c}$) model could be interpreted as favoring selection the 3-group full model (LR = 10.095, df = 5, P = 0.07). Among the 4 models, the AIC for the 3-group full model is smallest. These

Table. 2.

Estimates and measures of fit for adult meadow voles.

model	3-group full	3-group $\mathbf{p = c}$	2-group full	2-group $\mathbf{p = c}$
estimates π_1	0.704	0.605	0.729	0.703
p_1	0.117	0.297	0.312	0.334
c_1	0.493		0.353	
π_2	0.158	0.048	0.271	0.297
p_2	0.883	0.940	0.850	0.765
c_2	0.775		0.755	
$1 - \pi_1 - \pi_2$	0.138	0.347		
p_3	1.00	0.664		
c_3	0.223			
GOF test	6.62	16.35	16.44	16.41
df	11	14	14	16
P	0.8	0.3	0.3	0.4
L	-202.5729	-207.5876	-207.4876	-207.6204
AIC	421.146	425.175	424.975	421.241
$\hat{N}$	133	93	93	91
se	30	4.0	6.9	4.4
95%CI	(83, 193)	(85, 101)	(83, 107)	(83, 100)

statistics suggest that the 3-group full model is best.

In addition to the model comparison statistics, we have plots of fit for each model (Figures 3-1, 3-2, 3-3, and 3-4). These plots help in developing inferences on how the models work. Recaptures are modeled as having the same mixture as first captures. Plots of expected recapture frequencies by group illustrate how recapture data contribute to estimation of mixing parameters. Plots of expected frequencies of first capture by group show how these estimates relate to estimation of N. In the context of the biological knowledge about the species, these inferences can help characterize the population and lend credence to model selection. Plots for models that have poor fit may show that a restriction may be inappropriate or that the model needs more parameters.

In Figure 3.1a, 48 voles (cohort 1) that were first caught on occasion 1 have recaptures represented in Figure 3.1b. Dashed lines in Figure 3.1a show that groups 2 and 3 of the 3-group full model have similar first capture probabilities. However, in Figure 3.1b they have different recapture probabilities (dashed lines strongly differ). Figure 3.1 provides a graphic depiction

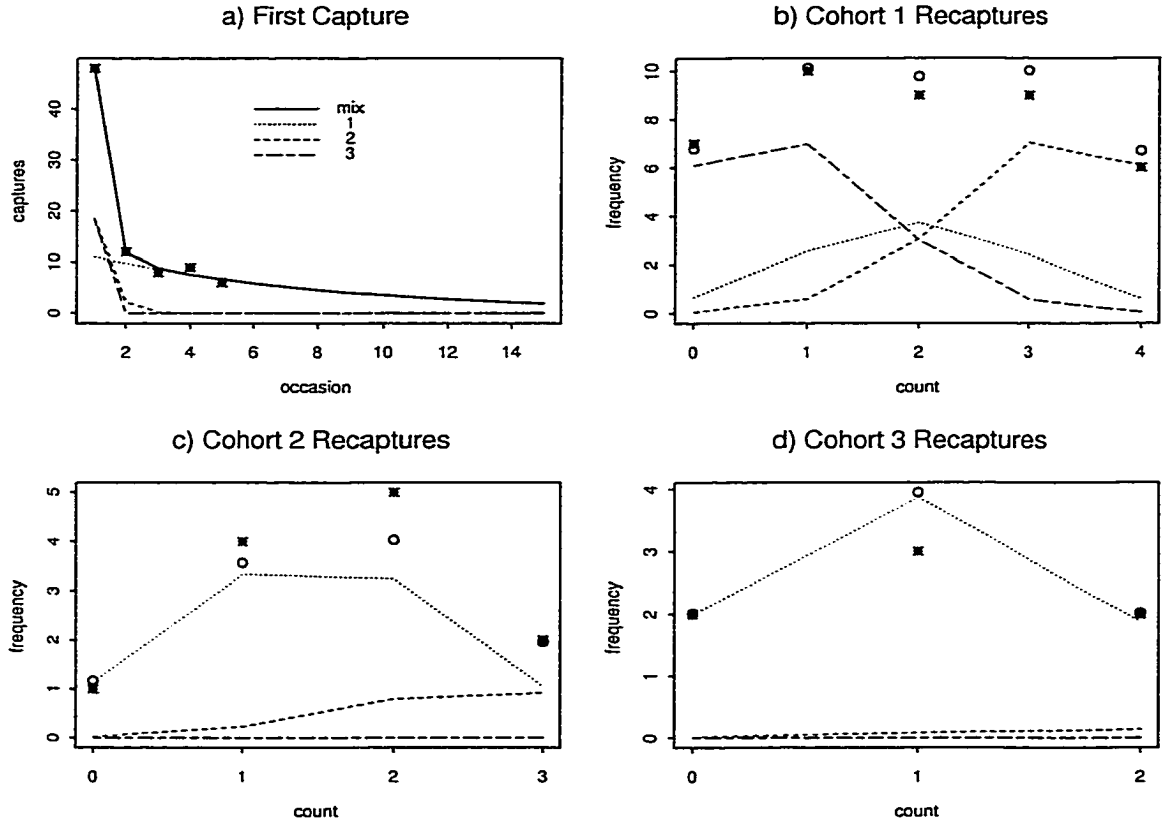


Figure 3.1 Observed and expected captures and frequencies of capture of meadow voles estimated with the 3-group full model. a) represents the fit of the mixture model (solid line), subdivided by the 3groups, to the time of first capture data. b), c), and d) represent recapture frequencies of cohort 1, 2, and 3 respectively. Asterisks are the observed frequencies, dash lines are expected group frequencies, and circles are expected mixture frequencies.

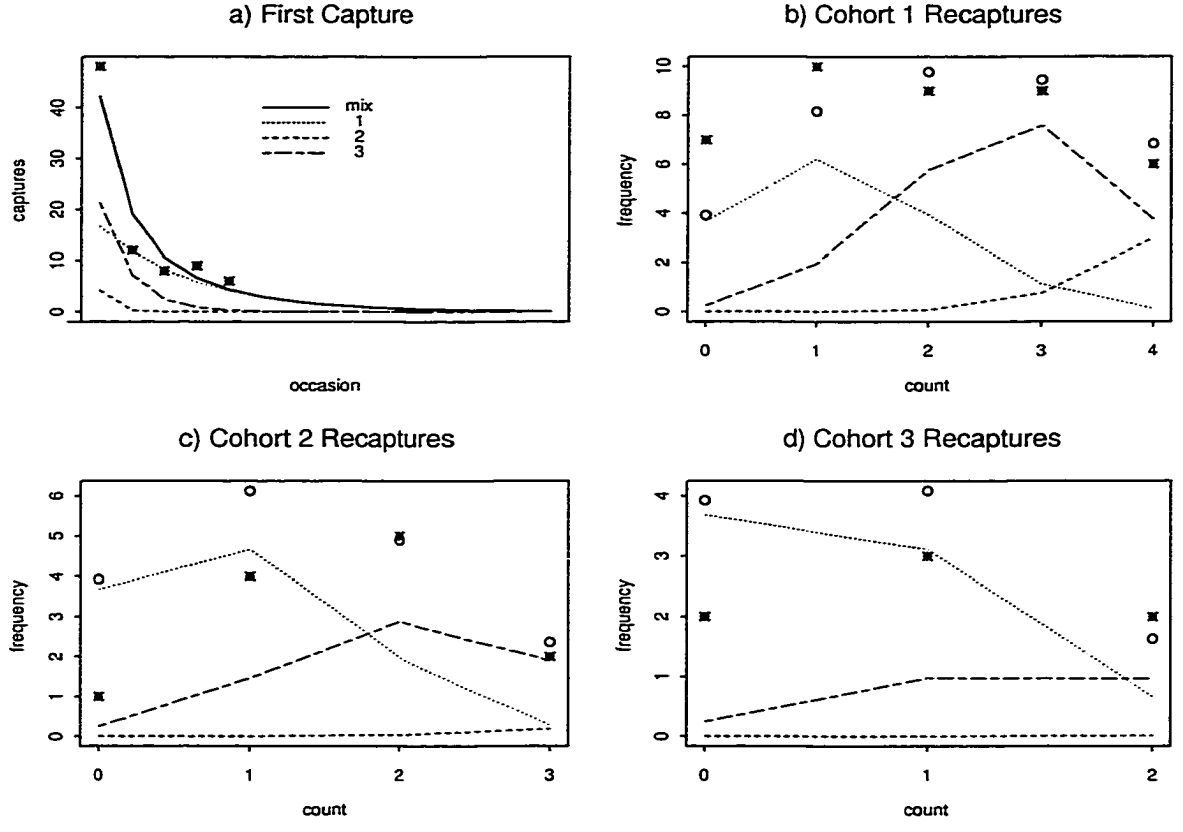


Figure 3.2 Observed and expected captures and frequencies of capture of meadow voles estimated with the 3-group restricted model ($\mathbf{p} = \mathbf{c}$). a) represents the fit of the mixture model (solid line), subdivided by the 3groups, to the time of first capture data. b), c), and d) represent recapture frequencies of cohort 1, 2, and 3 respectively. Asterisks are the observed frequencies, dash lines are expected group frequencies, and circles are expected mixture frequencies.

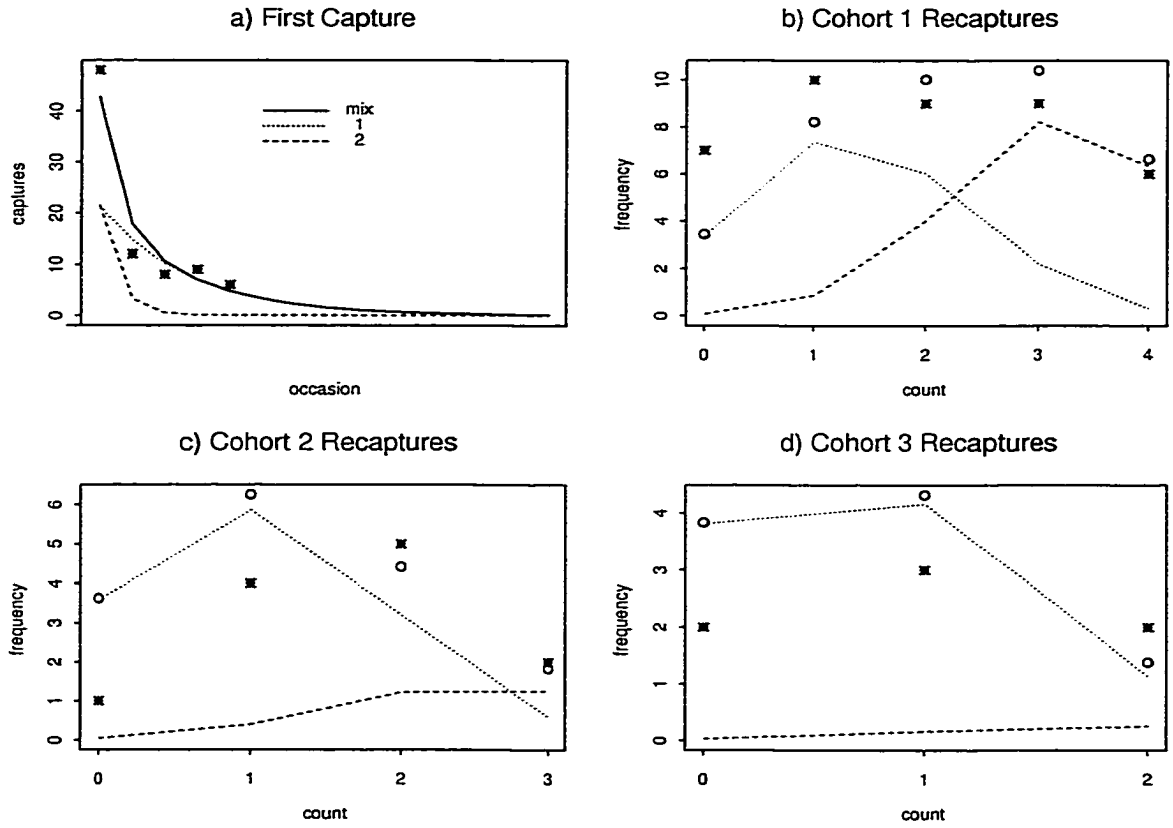


Figure 3.3 Observed and expected captures and frequencies of capture of meadow voles estimated with the 2-group full model. a) represents the fit of the mixture model (solid line), subdivided by the 3groups, to the time of first capture data. b), c), and d) represent recapture frequencies of cohort 1, 2, and 3 respectively. Asterisks are the observed frequencies, dash lines are expected group frequencies, and circles are expected mixture frequencies.

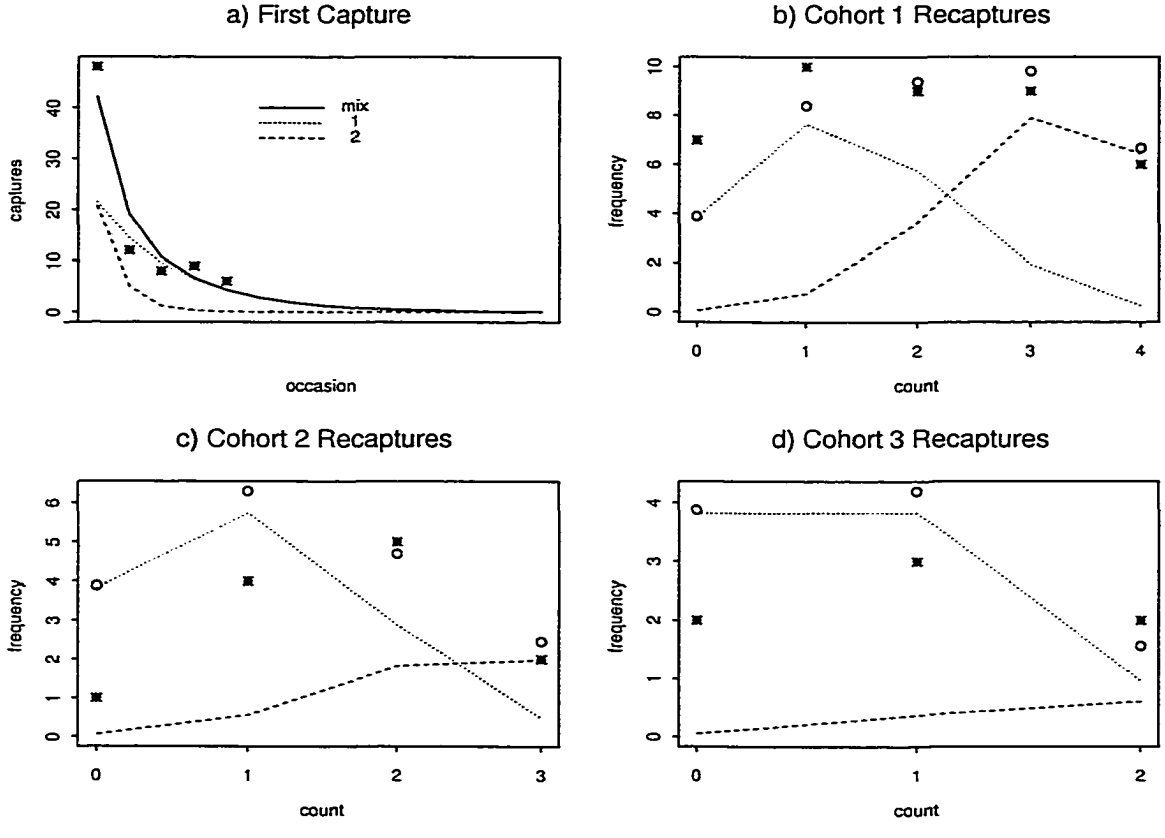


Figure 3.4 Observed and expected captures and frequencies of capture of meadow voles estimated with the 2-group restricted model ($\mathbf{p} = \mathbf{c}$). a) represents the fit of the mixture model (solid line), subdivided by the 3groups, to the time of first capture data. b), c), and d) represent recapture frequencies of cohort 1, 2, and 3 respectively. Asterisks are the observed frequencies, dash lines are expected group frequencies, and circles are expected mixture frequencies.

of how the 3-group full model uses recapture data to affect estimation of π's and p's through $\tau(\cdot)$. Among voles first caught on occasions 1, 2, and 3, the proportions of voles belonging to group 1 are $\tau_1(1) = 0.229$, $\tau_1(2) = 0.817$, and $\tau_1(3) = 0.971$, respectively. These values are reflected in Figure 3.1b by high frequency of recaptures for groups 2 and 3 in cohort 1, and in Figures 3-1c and 3-1d by low contributions to recaptures by groups 2 and 3 among cohorts 2 and 3.

The restriction $\mathbf{p} = \mathbf{c}$ has potential to provide better estimation of $\mathbf{p}$, but requires the assumption of no behavioral response to trapping. Estimation of $\mathbf{p}$ with the restriction $\mathbf{p} = \mathbf{c}$, as seen in Figures 3-2b, 3-2c, and 3-2d results in a fit to first capture data (Figure 3.2a) that is less satisfactory than the fit seen in Figure 3.1a for the unrestricted model. Consider the fit of the mixture and the way in which each group contributes to the expected frequencies in Figures 3-1b and 3-2b. In Figure 3.1b, the full model follows the pattern of observed recapture frequencies which may be the result of heterogeneity in behavioral response to trapping. The restricted model in Figure 3.2b does not fit the data as well as the full model.

The fit of the 3-group restricted model (Figure 3.2) is very similar to the fits of the full and restricted 2-group models (Figures 3-3 and 3-4). Comparisons between the full and restricted 2-group models suggest that allowing a behavioral response is of little advantage. On the contrary, small mammals often have a behavioral response to trapping and by allowing exposure to bait before trapping, the sampling design may have contributed to a behavioral response. Among the 4 models, the 3-group full model best explains the data.

In Table 2, the 3-group full model estimates of $\mathbf{p}$ and $\mathbf{c}$ show some heterogeneity in trap response. The recapture probability $\widehat{c}_1$ is greater than $\widehat{p}_1$, but $\widehat{c}_2$ is slightly smaller than $\widehat{p}_2$ and $\widehat{c}_3$ is substantially smaller than $\widehat{p}_3$. Group 1 might be described as "trap happy", while group 2 has no change in behavior. And group 3 could be thought of as "trap shy", but lack of novel stimuli on later trapping occasions is an alternative explanation (i.e., after the first time a vole investigates a trap, curiosity is not contributing to increase its recapture probability).

Traps were baited and held open for 2 days prior to sampling. We expect this period to affect probabilities of first capture much more than recapture probabilities. For example, it

is biologically reasonable that many voles have a moderate or low probability of first capture until they find bait in open traps, then switch to a higher probability of first capture before trapping starts. Other animals may not find the bait, thus have a lower probability of first capture when trapping starts. The baited period before trapping would not affect recapture probabilities, because all recaptures occur after a vole has been trapped at least once and is familiar with the bait and at least one trap. This scenario of responses to traps is consistent with the estimates of the 3-group full model.

For the 3-group model, $\widehat{N} = 133$ (se = 30) which translates to a naive density (133/1.1664ha) of 114 voles per ha. Although this is a high density, the mesic tall grass habitat at Doolittle Prairie is certainly capable of supporting this density (Getz et.al. 1987). In fact, voles at Doolittle were aggregated and no voles were captured in some areas with apparently good habitat.

3.3.2 Taxicabs

Carothers (1973) reported that 420 taxi cabs were registered in Edinburgh, Scotland during his mark-recapture study. This closed population was sampled for 10 consecutive days with observation points and times varied among days. Sighting a cab was considered a "capture". From Table 3, $\widehat{N}$ for the 2-group full model exactly matches the true population size. The MLE's of the 3-group full model, not given in Table 3, are $\widehat{\pi}_1 = 0.978$, $\widehat{p}_1 = 0.104$, $\widehat{c}_1 = 0.132$, $\widehat{\pi}_2 = 0.022$, $\widehat{p}_2 = 1$, $\widehat{c}_2 = 0.267$, $\widehat{\pi}_3 = 0$ and are equivalent to those of the 2-group full model (Figure 3.5, Table 3). Figure 3.5a shows that group 2 contributes to the expected number of first captures at the first occasion, but for occasions 2 through 10, all of the first captures come from group 1. In other words, this model suggests that probabilities of first capture among taxi cabs are homogeneous after occasion 1. In Figures 3-5c and 3-5d, probabilities of recapture are homogeneous. Perhaps the number of captures on the first occasion is the outcome of homogeneous probabilities of first capture.

The 2-group restricted model ($\mathbf{p} = \mathbf{c}$) resulted in estimates that are unreasonable (Table 3, Figure 3.6). In Figure 3.6a, the dotted line representing group 1 is flat and does not seem

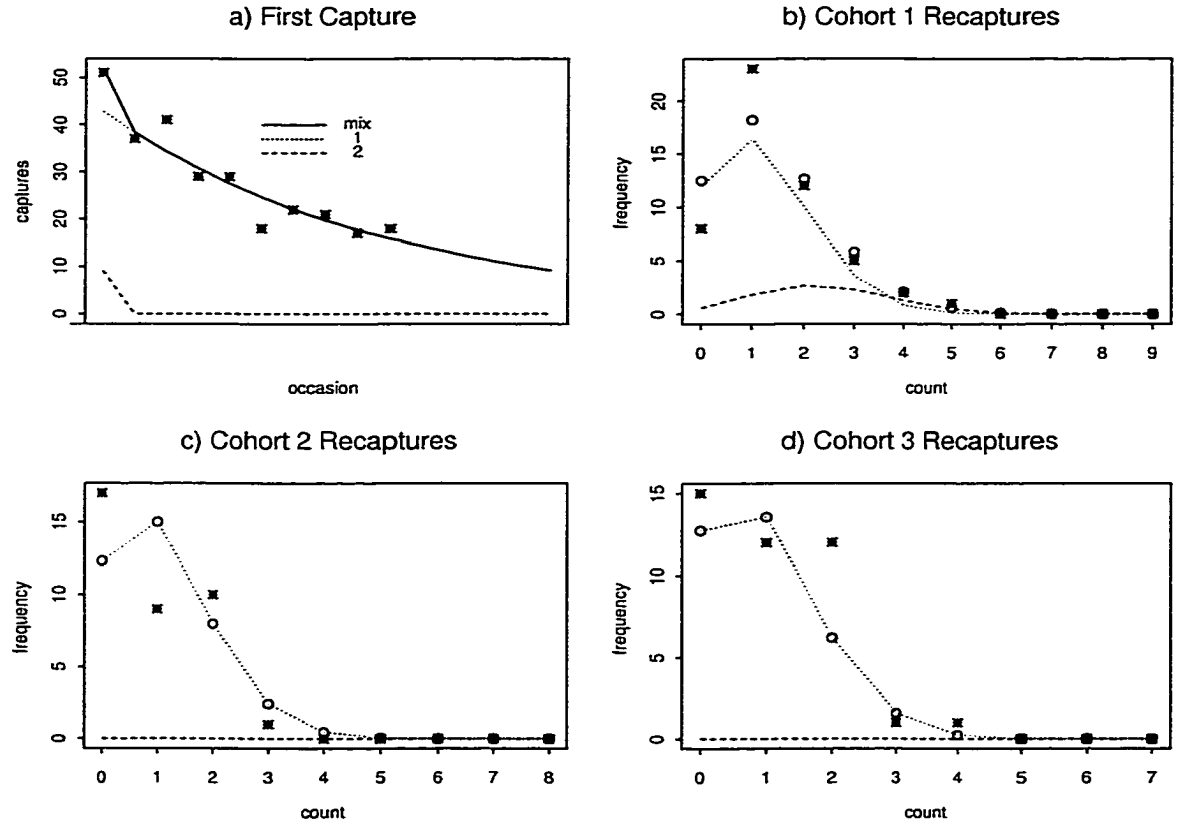


Figure 3.5 Observed and expected captures and frequencies of capture of taxi cabs estimated with the 2-group full model. a) represents the fit of the mixture model (solid line), subdivided by the 2 groups, to the time of first capture data. b), c), and d) represent recapture frequencies of cohort 1, 2, and 3 respectively. Asterisks are the observed frequencies, dash lines are expected group frequencies, and circles are expected mixture frequencies.

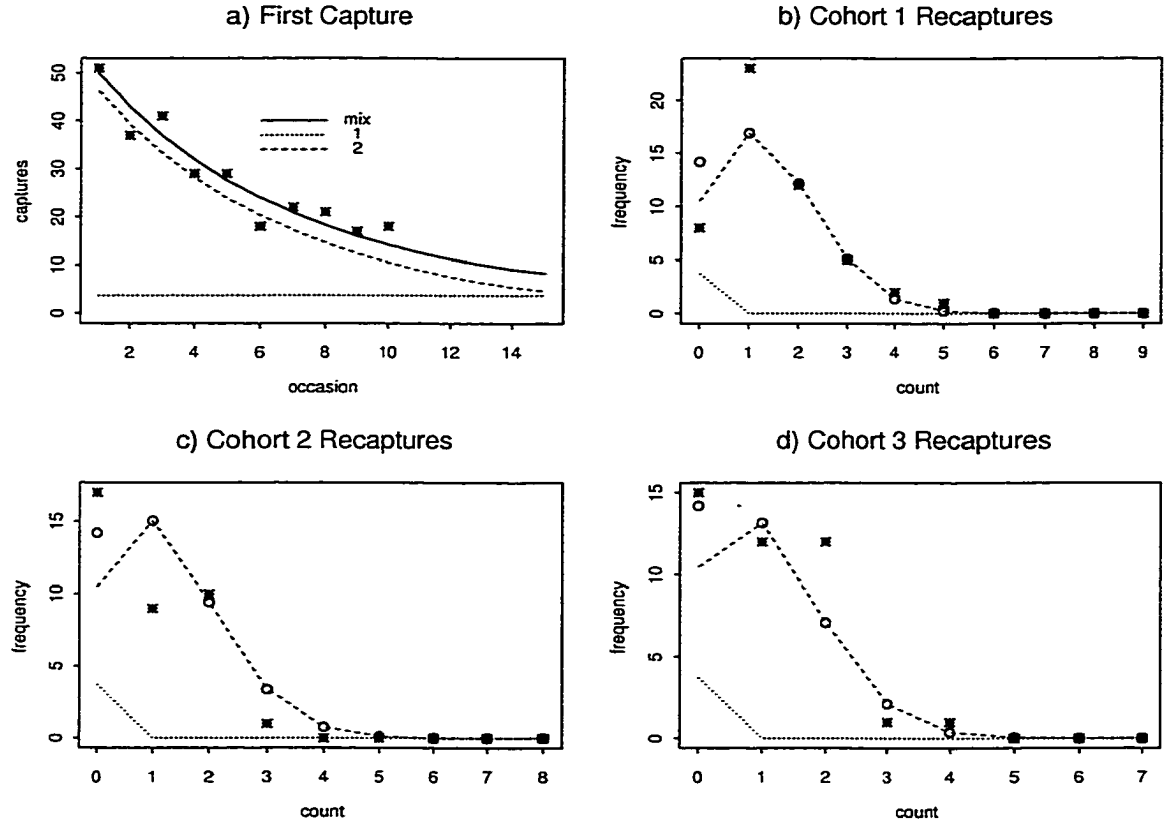


Figure 3.6 Observed and expected captures and frequencies of capture of taxi cabs estimated with the 2-group restricted model ($\mathbf{p} = \mathbf{c}$). a) represents the fit of the mixture model (solid line), subdivided by the 2 groups, to the time of first capture data. b), c), and d) represent recapture frequencies of cohort 1, 2, and 3 respectively. Asterisks are the observed frequencies, dash lines are expected group frequencies, and circles are expected mixture frequencies.

Table 3.
Estimates and measures of fit for taxicabs, Edinburgh Scotland.

models	2-group full	2-group $\mathbf{p = c}$	1-group full
estimates π_1	0.978	0.9999	1
p_1	0.104	0.000002	0.118
c_1	0.132		0.139
$1 - \pi_1$	0.023	0.0001	
p_2	1	0.152	
c_2	0.267		
GOF test	33.55	34.17	36.49
df	49	51	52
P	0.95	0.97	0.95
L	-927.333	-928.44	-928.98
AIC	1864.67	1862.88	1861.95
$\widehat{N}$	420	2400000	396
se	50		34
95%CI	(322, 518)		(329, 464)

to approach 0. This explains the offset of the mixture expectations from those of group 2 and the extremely large $\widehat{N}$ for this model. Projecting group 1 beyond occasion 10 and looking at the area under this curve suggests that an extremely large number of taxi cabs belong to group 1. Based only on size of the city, one could assume that Edinburgh has fewer than 10,000 taxi cabs, and the estimate of 2,400,000 is clearly extreme.

The 1-group model is not a finite mixture. This model is composed of a geometric model for the time of first capture and a binomial model for the number of recaptures. The estimate $\widehat{N}$ for the 1-group model is not a function of recapture parameters and is not influenced by recapture data in that an increase in recapture data does not change $\widehat{N}$ (Table 3, Figure 3.7). Estimates for the 1-group model are comparable to those of the homogeneous geometric removal model, $\widehat{N} = 396$ (se $= 34$), $\widehat{p} = 0.118$ (Chapter 2, section 2.1). By comparing Figures 3-5a and 3-7a, one can see the influence of occasion 1 on the estimate of $\widehat{p}$. Compared to the group 1 line in Figure 3.5a, the line for first captures in Figure 3.7a is "pulled" up on the left by the large observed number of captures on occasion 1. This steeper line corresponds to a greater probability of first capture, $\widehat{p} = 0.118$ in comparison to $\widehat{p}_2 = 0.104$ of the 2-group full model

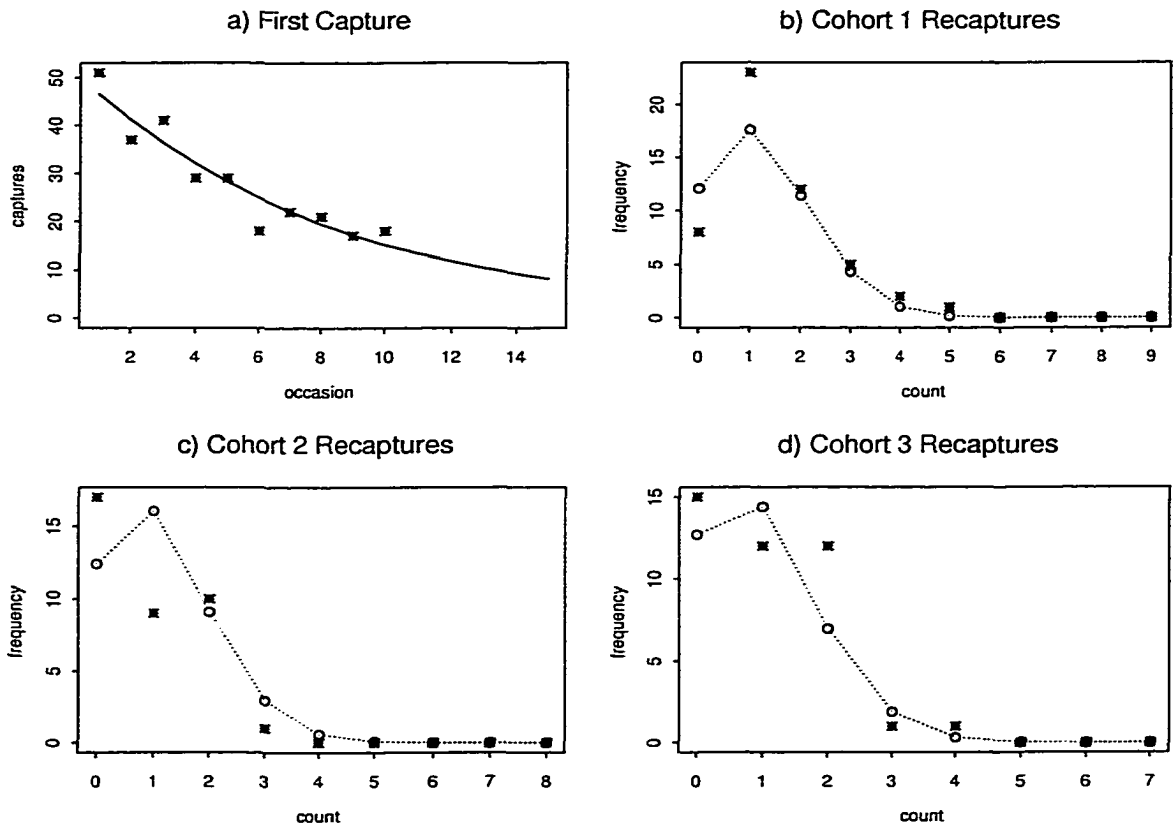


Figure 3.7 Observed and expected captures and frequencies of capture of taxi cabs estimated with the 1-group model. a) represents the fit of the model (solid line) to the time of first capture data. b), c), and d) represent recapture frequencies of cohort 1, 2, and 3 respectively. Asterisks are the observed frequencies, dash lines are expected group frequencies, and circles are expected mixture frequencies.

(Figure 3.5a). This is why $\widehat{N}$ under the 2-group full model is greater than $\widehat{N}$ under 1-group model.

From Figures 3-5, 3-6, and 3-7, and GOF statistics (Table 3), the fit is good for all three models. Comparisons between the 1- and 2-group models by LR and AIC suggest that the 1-group model is the best choice, though the point estimate for the 2-group model is most accurate. Selection of the 1-group model suggests that heterogeneity in capture probabilities among taxi cabs is not great.

In contrast to selection of the 1-group model, the model selection algorithm based on a multinomial framework in the program CAPTURE (Rextad and Burnham 1991) picks the heterogeneous model M_h and the interpolated jackknife estimate, $\widehat{N}_h = 471$ (se=27). Assumptions for the 1-group model are comparable to the multinomial behavioral model, M_b(Otis et.al. 1978). Two of these assumptions are the independence of capture events and constant probabilities of first capture and recapture through time and among taxi cabs. Estimates under M_b are $\widehat{N}_b = 393$ (se=35), $\widehat{p} = 0.119$, and $\widehat{c} = 0.139$. These M_b estimates are very similar to those of the 1-group model (Table 3). Model M_b was not highly ranked for selection by the program CAPTURE.

3.4 Discussion

The framework of finite mixtures for geometric time to first capture and binomial number of recaptures connects recapture data to estimation of the mixing distribution and probabilities of first capture through $\widehat{\tau}(x)$. The corresponding population size maximum likelihood estimator $\widehat{N}$ converges stochastically to N as the amount of data increases. This system models animals as belonging to groups as if each animal has latent probabilities for first capture and recapture that are constant within groups and across time. For the vole example, this group designation has a reasonable biological interpretation. Some animals have high first capture probabilities related to knowledge of bait and traps before trapping begins. Other animals with smaller first capture probabilities may be less mobile and do not encounter the traps until after trapping begins. Heterogeneity in capture probabilities may be animal-specific rather than a group

response. In this case, the estimated finite mixture is an approximation.

Burnham (1972) used a truncated beta-binomial mixture to model heterogeneous capture probabilities in the absence of a behavioral response to trapping. For this model, applied to a given data set, large estimates of N fit as well as small estimates. A continuous mixing distribution, such as a beta, has greater potential to put mass at p close to 0 than does a finite mixture. Some geometric finite mixture models and data have $\widehat{p}$'s near 0 and an $\widehat{N}$ that is unreasonably large (e.g., Section 3.2 taxi cab 2-group restricted model, $\mathbf{p} = \mathbf{c}$). Though this is a problem, it is not a problem with all data and models.

The geometric finite mixture framework offers a new model-based estimator for data with heterogeneity in recapture probabilities when a trapping behavioral response exists. This framework allows modeling of heterogeneity in behavioral response to trapping (e.g., $p_1 > c_1$ and $p_2 < c_2$ may occur in the same model). With this framework, modeling of recapture data can lead to better estimation. From the vole example in Chapter 2, the 2-group geometric removal model population estimate is $\widehat{N} = 129$ (se=113). In this chapter, the 3-group full model estimate is $\widehat{N} = 133$ (se=30). The latter model characterizes the animal capture activity more completely and offers greater precision in estimation of population size, N.

Do geometric finite mixture models provide better estmates than jackknife estimators or estimators based on multinomial models? At what levels of capture probabilities and heterogeneity do these new estimators have substantial bias or poor precision? How many occasions are needed to effectively model mark-recapture data with geometric finite mixtures? Simulations of mark-recapture studies can answer these questions and may suggest areas of further improvement.

Modeling mark-recapture data with geometric finite mixtures requires the assumption that probabilities of first capture ($\mathbf{p}$) and recapture ($\mathbf{c}$) are constant through all k occasions (i.e., through time). Environmental pertubations such as storm events or changes in temperature may affect the activity of animals and their probability of capture. Analysis of the influence of each occasion on estimates of θ and N may show when environmental changes may have affected the sampling efficiency and capture probabilities. In the taxi cab 2-group model, heterogeneity

occurs among taxicabs sighted on the first occasion ($\widehat{p}_2 = 1$, so all group 2 taxicabs were seen on occasion 1). Taxi cabs sighted on subsequent occasions have homogeneous probabilities of first "capture". The contrast between the 2-group full model (Table 3) and the 1-group model is very similar to an analysis of influence by the first occasion. By deleting from the data the 51 taxi cabs seen on the first occasion, we get deletion statistics of $\widehat{p} = 0.106$ and total $\widehat{N} = 364 + 51 = 415$ from a 1-group full model. The influence of occasion 1 is a change in $\widehat{p}$ of 0.012 and in $\widehat{N}$ of 19. This information can be combined with background information on the study to decide how to proceed with modeling. This example works for occasion 1, but the method needs to be generalized for measurement of influence at other occasions.

The homogeneous geometric distribution is similar to an interval-censored exponential distribution. Modeling first captures with an interval-censored Weibull distribution with shape parameter $\beta < 1$ would be an alternative to modeling with a geometric finite mixture. One advantage would be a reduction in the number of parameters required. Two major disadvantages are 1) the parameters do not have a direct interpretation relating to animal capture activity and 2) the connection of recapture data to estimation of Weibull parameters is missing. Modeling of first capture and recapture data by finite mixtures allows a connection between recapture data and estimation of N and uses model parameters with direct interpretations related to animal captures.

CHAPTER 4. SIMULATION

4.1 Introduction

In Chapters 2 and 3 we model animals as belonging to discrete groups in which probabilities of first capture and recapture are constant. In these previous chapters, we applied estimation with our methods to example ecological data, but the true values of parameters were seldom known. Herein, we simulate the capture process on a population of known size (N) with each animal having known group affiliation and known capture probabilities. For a given set of mixture proportion and capture probability parameters, we simulate the capture process to create many replicate sets of data. We use estimation of N from these simulated data to study the bias, precision, and other operating characteristics of our estimators in context of known parameter values and underlying assumptions.

We have organized the simulation study as three sections in which parameters and analysis are chosen to answer different questions about the new estimators. How well do our geometric finite mixture estimators perform compared to the Pollock and Otto (1983) jackknife estimator and the Burnham (1972) jackknife estimator that are most frequently used to estimate population size under heterogeneous capture probabilities? The first section tries to answer this question with varying capture probabilities, heterogeneity, and number of occasions, but with constant $N = 200$ and no behavioral response. The second section simulates homogeneous populations to examine the effects of small capture probabilities and the number of occasions on distribution of $\widehat{N}$'s without confounding with heterogeneity. Also, we look at how readily heterogeneity is incorrectly indicated in our model selection methods. In the third section we address how well our estimators describe the underlying capture process and how they perform with behavioral response to first capture. In these three sections we describe results with

estimates of bias, root mean squared error (RMSE), and frequency plots of estimates.

4.2 Methods

For each simulation, data generation started with v groups of N animals (e.g., a total of 400 animals for $v = 2$ and $N = 200$). Each simulated animal was assigned probability of first capture p according to group. These assignments were constant through all replicates. For each animal and replicate, a capture history was created by a Bernoulli trial with p as the probability of "success" (capture) at each of 20 occasions. All animals were modeled to survive through the 20 occasions (i.e., no accidental death). From these vN capture histories on each replicate, the first π_l animals of group l for $l = 1, \ldots, v$ were used for a total of N animals. The purpose for generating the vN capture histories was to allow inferences about effects of changes in p and k from 100 replicates though average estimates had not converged. Later, these simulations were extended to at least 500 replicates. With average estimates closer to convergence, the design described above is less important to the inferences. For each replicate, results of the first k trials were compiled across animals to form data sets for $k = 5$ and 10. These values of k are consistent with example data sets in Chapters 2 and 3. When $k = 20$, the number of animals seen in a study is much closer to N and the success of modeling is greater, however studies of this length are not typical in practice. The structure of these data sets were similar to those in Chapters 2 and 3. Estimation was by the direct search algorithm of Hooke and Jeeves (1961) with the true parameter values as starting points. The searches were limited to 500 steps.

In an animal study, other research or literature gives the investigator an upper bound for what N is reasonably possible. Estimates that are significantly greater than this bound can be ruled out. In the Chapter 1 pocket mouse example, the P2 estimate is 170000. This estimate is too large. Norris and Pollock (1996) use $2N$ as an upper limit above which replicate estimates are excluded from a bootstrap analysis. We use $5N$ as an upperlimit for excluding estimates from simulation statistics.

In simulation studies, Otis et.al. (1978) and Burnham and Overton (1979) use the number

of replicate 95% confidence intervals that include N (coverage) as a simulation statistic. This is a function of the point estimate and the estimate of its variance. An estimator of N can be accurate and precise, but if the associated estimator of variance is poor, then coverage by the 95% confidence interval may falsely indicate poor estimation by the point estimator. Such a point estimator could have a poor coverage rate. Evaluation of an estimator through coverage depends on asymptotic normality of the estimator, though at small p and k the distribution may be skewed. We use the root mean squared error,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(\widehat{N}_i - N\right)^2}{n-1}}$$

as a simulation sample statistic. For a set of replicated data, alternative estimators can be compared by their RMSE. A smaller RMSE indicates that the estimator tends to provide estimates closer to the true value of the parameter.

4.3 Comparison to jackknife estimators

Chapters 2 and 3 provide development of estimators of population size using geometric finite mixtures. The objective of this section is to determine conditions of capture probability and heterogeneity under which these new estimators perform well. A second objective is to compare the new estimators to other estimators. In Chapter 2, we modeled the process of first capture on a population of size N with a 2 component geometric finite mixture (P2). In Chapter 2, one of our models of first capture and recapture with no behavioral response is a 2 component finite mixture of geometric first capture and binomial recapture (PeqC2). We assumed that animals belong to groups in which the probability of first capture (p) and probability of recapture (c) are constant. When capture probabilities are heterogeneous among animals, commonly used methods for estimating N are Pollock and Otto's (1983) jackknife estimator ($\widehat{N}_{Jb}$) for $p \neq c$ and Burnham's (1972) jackknife estimator ($\widehat{N}_J$) for $p = c$. We compare our estimators to these 2 jackknife estimators using the ratio of RMSE as a measure of relative efficiency. In our simulations, π is the proportion of animals in group 1, p_1 and p_2 are probabilities of first capture for groups 1 and 2, respectively, and k is the number of

sampling occasions. We simulate the capture process on $N = 200$ animals at each combination of $\pi = 0.5, 0.7, 0.9$; $p_1 = 0.1, 0.2$; $p_2 = 0.5, 0.9$; and $k = 5, 10$. Simulations with the same values for p_1 and p_2 used the same capture histories, but had different proportions from each group according to π. For instance, with $p_1 = 0.1$ and $p_2 = 0.5$, the simulations at $\pi = 0.5$ and 0.7 share 80% of their capture histories. In each replicate of 200 animals at $\pi = 0.5$, 100 capture histories come from each group; at $\pi = 0.7$, 40 more group-1 capture histories and 40 fewer group-2 capture histories are used. This overlap of simulations facilitates contrasts between levels of heterogeneity when 100 replicates are used and average estimates have not converged, but it is of little consequence when average estimates are close to convergence.

4.3.1 Removal

The 2-group finite mixture model (P2) RMSE's were estimated by simulation for all combinations of $\pi = 0.5, 0.7, 0.9$, $p_1 = 0.1, 0.2$, $p_2 = 0.5, 0.9$, and $k = 5, 10$ (Appendix C Table C1). Comparisons on each horizontal line in Figure 4-1 show that estimated RMSE(P2)'s are greater at $k = 5$ than at $k = 10$ for all simulated combinations of π, p_1, and p_2. The RMSE(P2)'s are smaller for data sets with more occasions, because the model P2 maximum likelihood estimate (MLE) of N is statistically consistent (Bain and Engelhardt 1989).

Among simulations with $p_2 = 0.9$ in the top half of Figure 4-1, RMSE(P2)'s are smaller with $p_1 = 0.2$ than with $p_1 = 0.1$. Among simulations with $p_2 = 0.5$, RMSE(P2)'s do not substantially differ between $p_1 = 0.1$ and $p_1 = 0.2$. Perhaps with $p_2 = 0.5$, the model produces data where the effect of heterogeneity is more subtle and difficult to fit than when $p_2 = 0.9$.

The RMSE(P2) estimates are likely to be smaller than the true RMSE(P2)'s, because we did not use estimates greater than $5N$ (1000) which were obviously too large. In some replicates, estimates were not found because the direct search of the MLE was limited to 500 steps. Model P2 has 3 parameters and estimation of these from 5 occasions of data is difficult. None of the 24 simulations had model P2 estimates for more than 80% of the replicates. By contrast, the closed form $\widehat{N}_{Jb}$ estimator had an estimate for every replicate.

Exact RMSE's for $\widehat{N}_{Jb}$ were calculated from the mean and variance according to Pollock

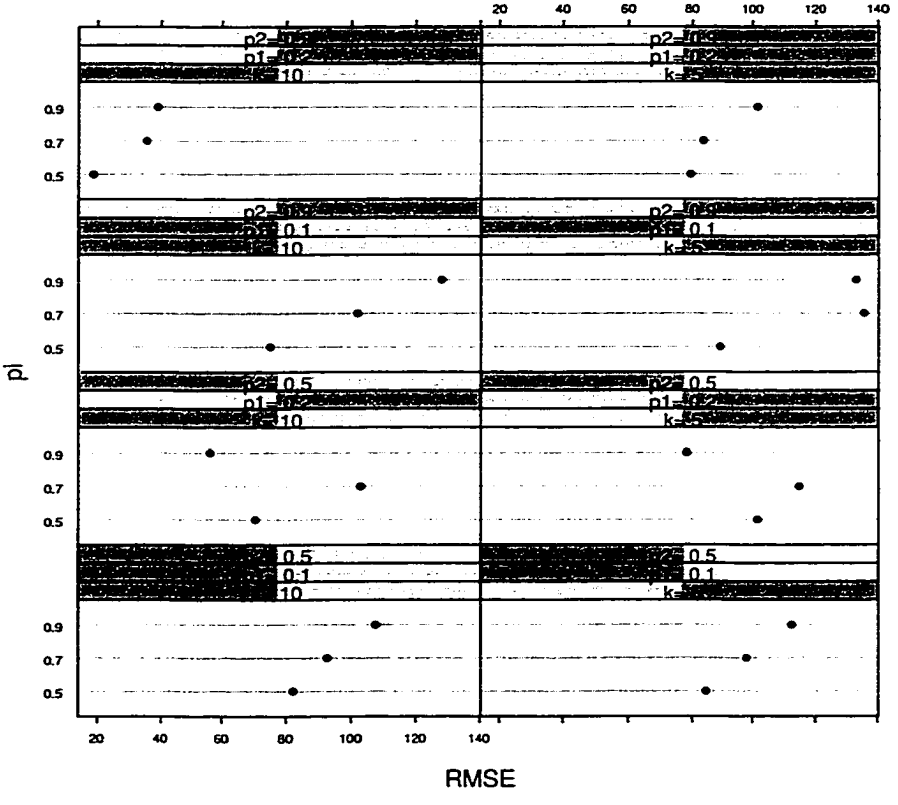


Figure 4.1 RMSE(P2) at $k = 5$ and 10, $p_1 = 0.1$ and 0.2, $p_2 = 0.5$ and 0.9, and $\pi = 0.5$, 0.7, and 0.9.

and Otto 1983 equations (12) and (13) (Appendix C, Table C2). In the top half of Figure 4-2, the comparisons of pairs of RMSE$\left(\widehat{N}_{Jb}\right)$'s on each horizontal line show that for $p_2 = 0.9$, RMSE$\left(\widehat{N}_{Jb}\right)$'s are smaller when $k = 10$. However, for $p_2 = 0.5$ in the lower half of Figure 4-2, RMSE$\left(\widehat{N}_{Jb}\right)$'s are greater when $k = 10$ compared to $k = 5$. This result means that when $p_2 = 0.5$, additional 5 occasions do not provide estimates closer to N. This characteristic can be seen in more detail in Figure 4-3, where RMSE$(\widehat{N}_{Jb})$ does not monotonically approach 0 as k increases. For animal surveys where the smallest probability of first capture among animals is $p = 0.1$, increasing k from 5 to 10 improves characteristics of $\widehat{N}_{Jb}$ (i.e., reduces RMSE), however if this smallest p is 0.2, then the same increase in sampling does not improve $\widehat{N}_{Jb}$. This pattern holds for $\pi = 0.5, 0.7, 0.9$.

Efficiency of the P2 estimator is measured with the ratio of the exact RMSE$(\widehat{N}_{Jb})$ divided by the estimated RMSE(P2). When this ratio is greater than one, the model P2 has a better estimator of N. In Figure 4-4, relative efficiency is only greater than 1 for one parameter combination, $\theta' = (\pi, p_1, p_2) = (.5, .2, .9), k = 10$. Otherwise, relative efficiency is less than 1, which suggests that the Pollock and Otto (1983) jackknife, $\widehat{N}_{Jb}$ is a better estimator for removal data with heterogeneous probability of first capture among animals.

4.3.2 Mark-recapture

We estimated the reduced ($\mathbf{p = c}$) 2-group finite mixture mark-recapture model (PeqC2) RMSE's by simulation for all combinations of $\pi = 0.5, 0.7, 0.9$, $p_1 = 0.1, 0.2$, $p_2 = 0.5, 0.9$, and $k = 5, 10$ (Appendix C Table C3). For each of these simulations, we also estimated RMSE's for the Burnham and Overton (1978) jackknife estimator which requires the same assumptions as PeqC2. These assumptions are 1) independence of capture events, 2) constant capture probabilities through time, 3) no behavioral effect of first capture, and 4) population closure. In mark-recapture studies, as occasions increase or as probabilities of first capture increase, the marked portion of the population (i.e., the amount of data) increases. Ideally, RMSE's get smaller as data increase.

Comparison of RMSE(PeqC2)'s on each horizontal line of Figure 4-5 shows that RMSE(PeqC2)'s

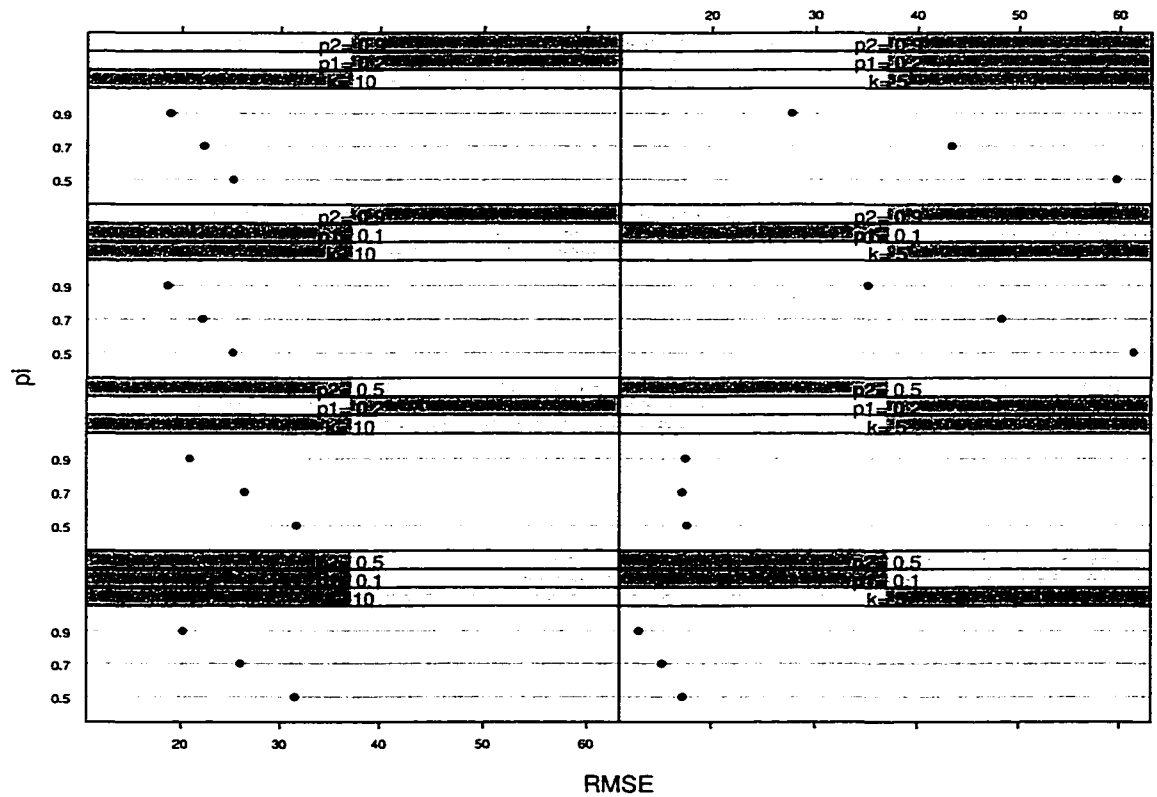


Figure 4.2 Exact RMSE($\widehat{N}_{Jb}$) at $k = 5$ and 10, $p_1 = 0.1$ and 0.2, $p_2 = 0.5$ and 0.9, and $\pi = 0.5$, 0.7, and 0.9.

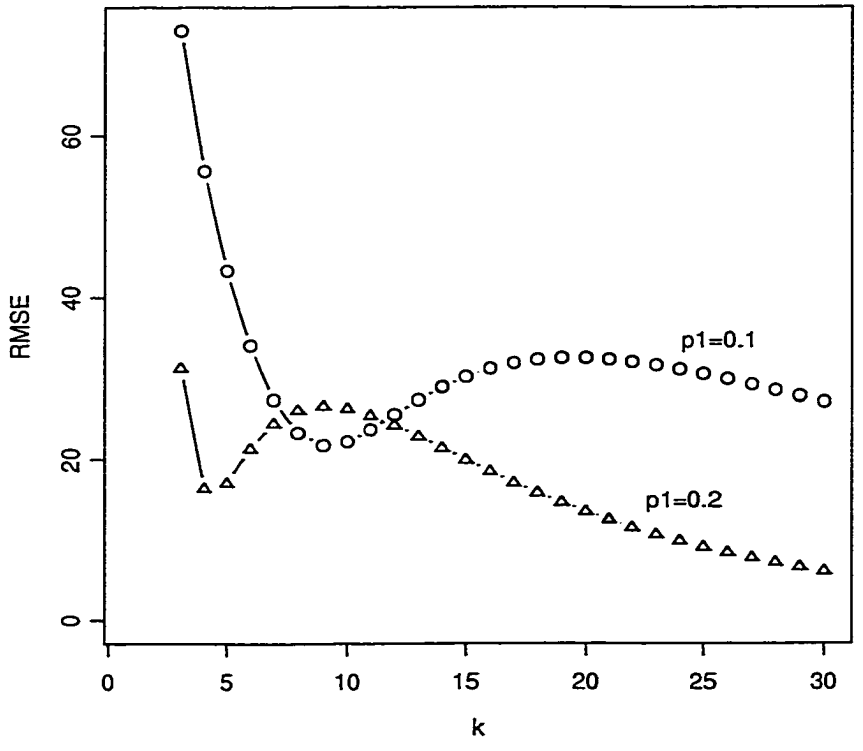


Figure 4.3 Exact RMSE($\widehat{N}_{Jb}$) at $p_2 = 0.5$ and $\pi = 0.7$ for $p_1 = 0.1$ (circle), 0.2 (triangle) and $k = 3, \ldots, 30$.

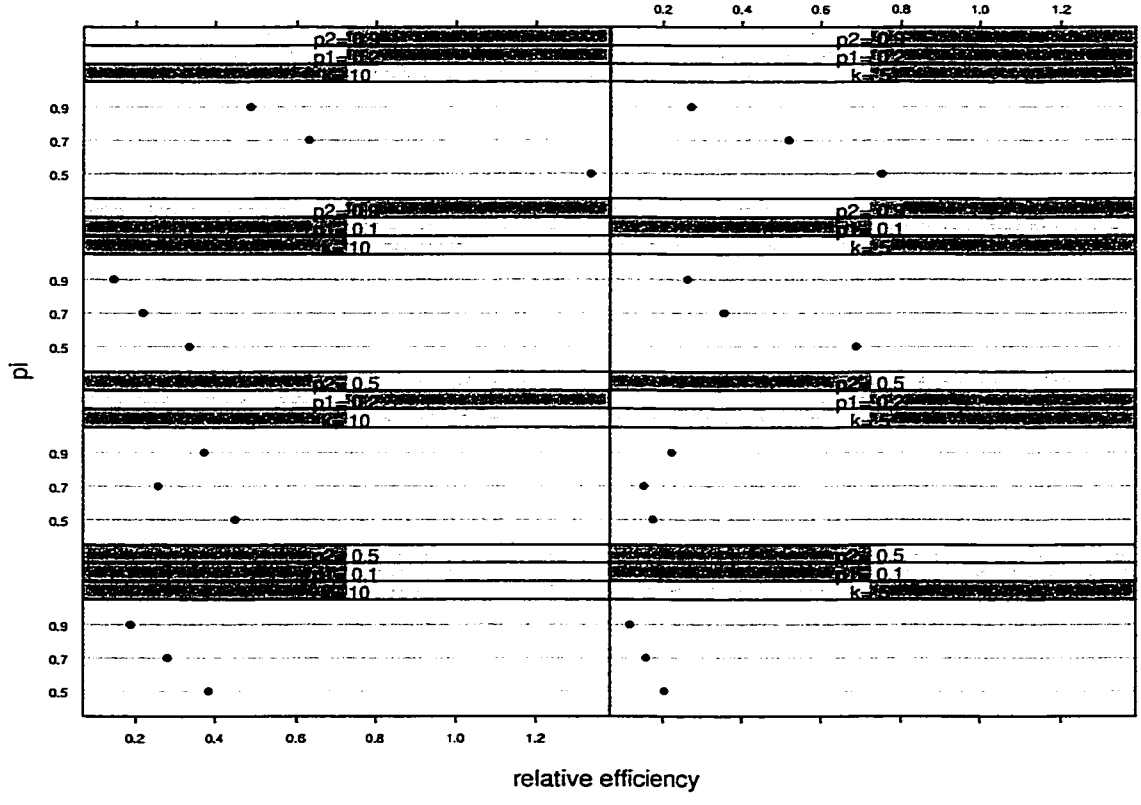


Figure 4.4 Relative efficiency = RMSE($\widehat{N}_{Jb}$)/RMSE(P2) at $k = 5$ and 10, $p_1 = 0.1$ and 0.2, $p_2 = 0.5$ and 0.9, and $\pi = 0.5$, 0.7, and 0.9. Relative efficiency greater than 1 indicates that the 2-group geometric finite mixture estimator is a better estimator of N than the Pollock and Otto (1983) jackknife estimator.

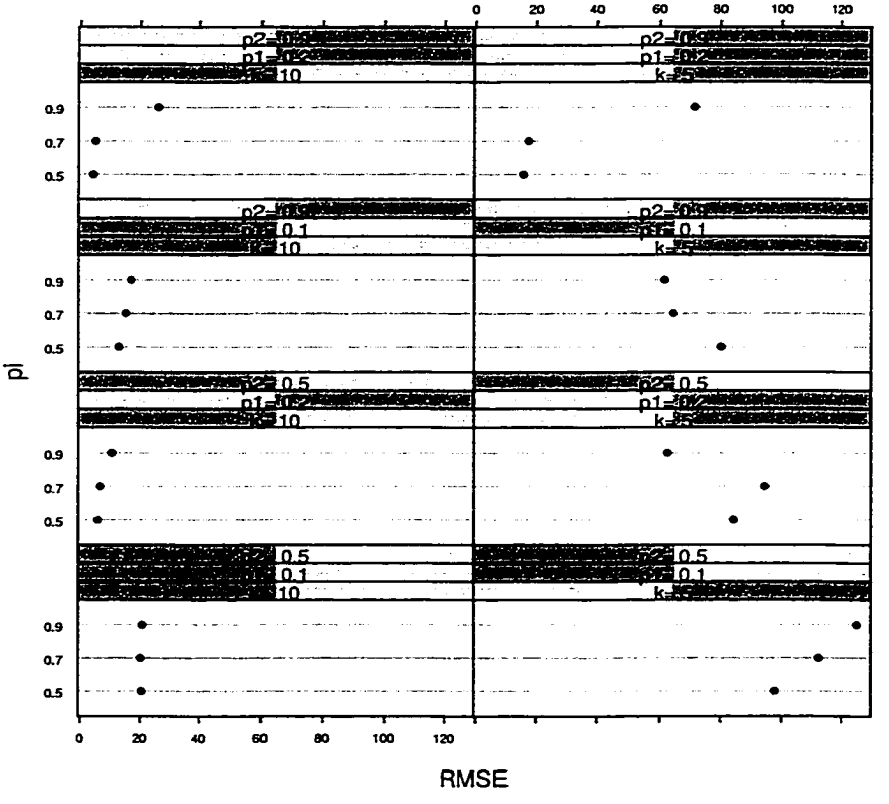


Figure 4.5 RMSE(PeqC2) at $k = 5$ and 10, $p_1 = 0.1$ and 0.2, $p_2 = 0.5$ and 0.9, and $\pi = 0.5$, 0.7, and 0.9.

are smaller at $k = 10$. At $\pi = 0.9$, $p_2 = 0.9$, and $k = 5$ (lines 1 and 4 in the upper right corner of Figure 4-5), the RMSE(PeqC2) is smaller for $p_1 = 0.1$ than for $p_1 = 0.2$, though not substantially. Otherwise, for all other simulations, RMSE(PeqC2)'s are smaller at $p_1 = 0.2$ (e.g., in the lower left corner of Figure 4-5, at $p_2 = 0.5$ and $k = 10$). These patterns show that with increase in the smallest probability of first capture or a greater number of occasions the PeqC2 model estimator provides estimates closer to N.

By contrast, for the Burham and Overton (1978) jackknife estimator $\left(\widehat{N}_J\right)$ at $k = 10$ (left side of Figure 4-6) the RMSE's are smaller when $p_1 = 0.2$ compared to when $p_1 = 0.1$, which is what you would hope to see. However, for $k = 5$ (right side of Figure 4-6) the greater p_1 does not provide smaller RMSE($\widehat{N}_J$)'s. Also, comparisons on each horizontal line of Figure 4-6 show that at $p_1 = 0.1$ RMSE($\widehat{N}_J$)'s are larger for $k = 10$ than for $k = 5$. Increases in sampling may not produce $\widehat{N}_J$ estimates closer to N. Figure 4-7 (left side) shows that for $k = 10$, the PeqC2 estimator is more efficient than $\widehat{N}_J$. For $k = 5$ (right side), the PeqC2 estimator is only better at $p_1 = 0.2$, $\pi = 0.7, 0.5$ (lines 2 and 3 from the top).

4.4 Homogeneous Population

In designing mark-recapture studies, the researcher must consider how much trapping effort to use and how long to run the survey. Survey design for these concerns can impact the validity of assumptions (e.g., closure) and it can affect the amount of information for estimation. With fewer members of a population observed, estimation of N through a homogeneous geometric model is more difficult. To investigate this problem, we simulated homogeneous populations with $p = 0.1$ and 0.2. The homogeneous geometric model is basic to removal methods developed in Chapter 2. By investigating performance for homogeneous populations, we can suggest how much of the performance of the geometric finite mixture model estimators is due to heterogeneity as opposed to a low capture probability or a short survey.

In the simulation with $k = 5$, $p = 0.1$, and $N = 200$, we were able to produce estimates for 859 of 1000 replicates for the homogeneous geometric removal model (P) (Appendix D Table D1). When the sampling occasions were extended to $k = 10$, we produced estimates

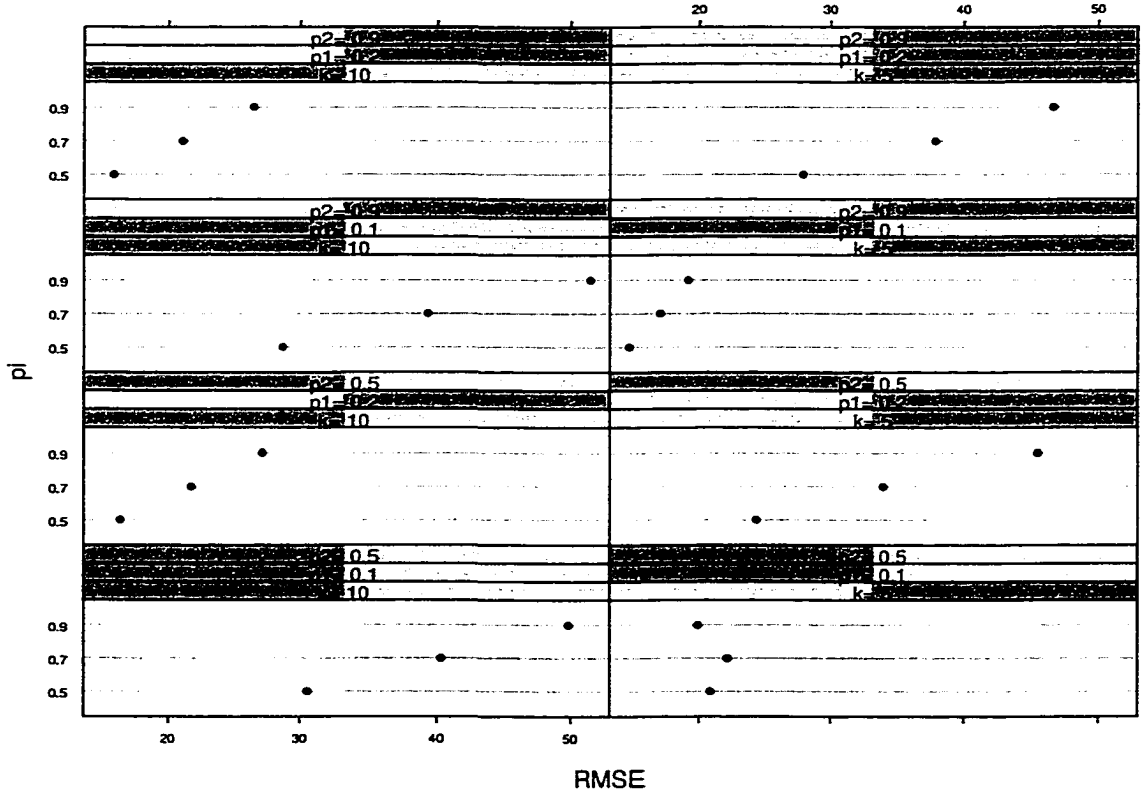
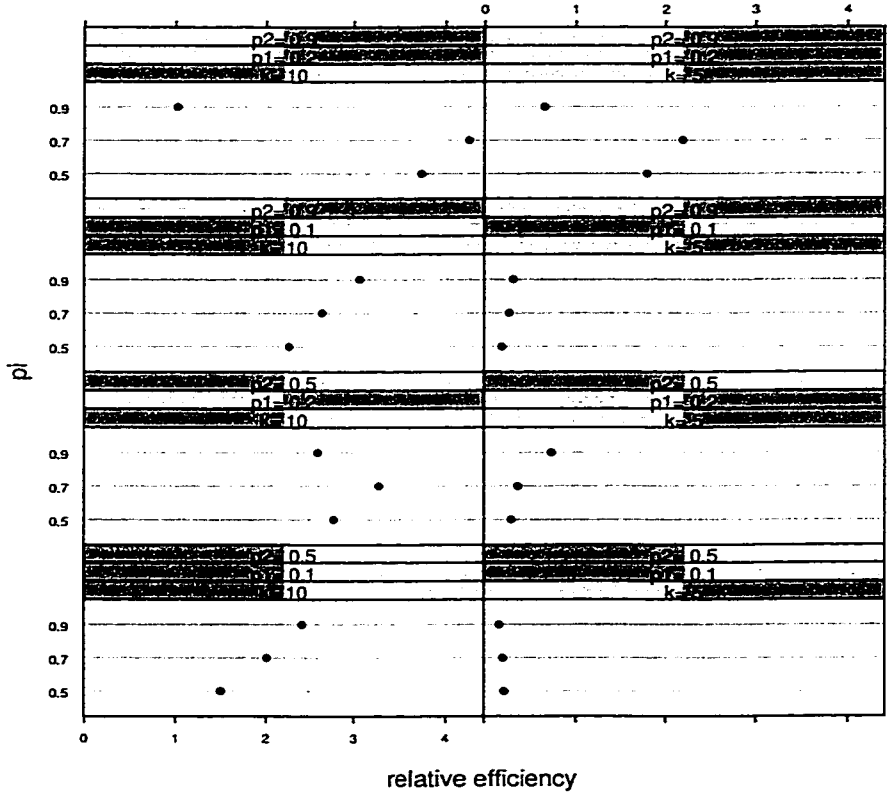


Figure 4.6 RMSE(jack) at $k = 5$ and 10, $p_1 = 0.1$ and 0.2, $p_2 = 0.5$ and 0.9, and $\pi = 0.5$, 0.7, and 0.9.


0
1
2
3
4
p2=0.5
p1=0.2
k=10
0.9
0.7
0.5
p2=0.9
p1=0.1
k=10
p2=0.5
p1=0.1
k=5
p2=0.5
p1=0.2
k=5
p2=0.5
p1=0.1
k=10
p2=0.5
p1=0.1
k=5
p2=0.5
p1=0.2
k=10
p2=0.5

for all replicates. Also, in the simulation with $p = 0.2$ (Appendix D Table D3), there were estimates for all replicates. With larger k or p, more animals are observed on average and such data provide for better estimation of N through model P. Figure 4-8 shows data of the first 16 replicates of the simulation with $p = 0.1$ and $k = 5$. The line indicates expected values under the model that created the data. Circles represent frequencies that randomly occurred. The observed frequencies of replicate number 13 neatly follow the expected frequencies under the model. By contrast, in replicates 1, 5, 10, 11, and 14 the data do not show a decline in the frequencies of first captures through time. With little or no negative slope to the line through the frequencies, the inference is that the capture probability is very close to zero. The corresponding log likelihood curves are displayed in Figure 4-9 with a dotted vertical line showing the MLE for p very close to the edge of the parameter space. Replicates 1, 5, 10, 11, and 14 have extremely large $\widehat{N} = 2.8 \times 10^9$, 2.7×10^9, 894, 1125, and 1073, respectively, but the true N is 200. Low estimates can also occur. In Figure 4-8, replicates 4, 12, and 16 show a fast decline in frequencies. These result in low estimates, $\widehat{N} = 114$, 117, and 121 for replicates 4, 12, and 16, respectively. Figure 4-10 shows the first 16 replicates extended to $k = 10$ and Figure 4-11 shows that maximums of the log likelihood curves are well away from the edge of the parameter space, away from 0. For these data, the model P performs better with $\widehat{N}$'s ranging from 156 to 263. These simulations indicate that when a low proportion of the population is observed during a survey, patterns of data can occur that make estimation difficult even though no model assumptions are violated.

For homogeneous capture probabilities, the Zippin (1956) estimator is a commonly used removal estimator. Figure 4-12 shows the correlation between the Zippin estimates and the model P estimates. Data that are difficult for estimation with model P are also difficult with the Zippin estimator. For $p = 0.1$, model P estimates tend to be larger than the corresponding Zippin estimates (Figure 4-12a). For $p = 0.2$, the estimates are closer (Figure 4-12b).

We investigated the distribution of $\widehat{N}$ for model P. From simulations of removal from homogeneous populations, when $p = 0.1$ and $N = 200$, $\widehat{N}$ is right skewed (Figures 4-13 a and b), though less so for $k = 10$ than $k = 5$. Some $\widehat{p}$'s come close to the edge of the parameter

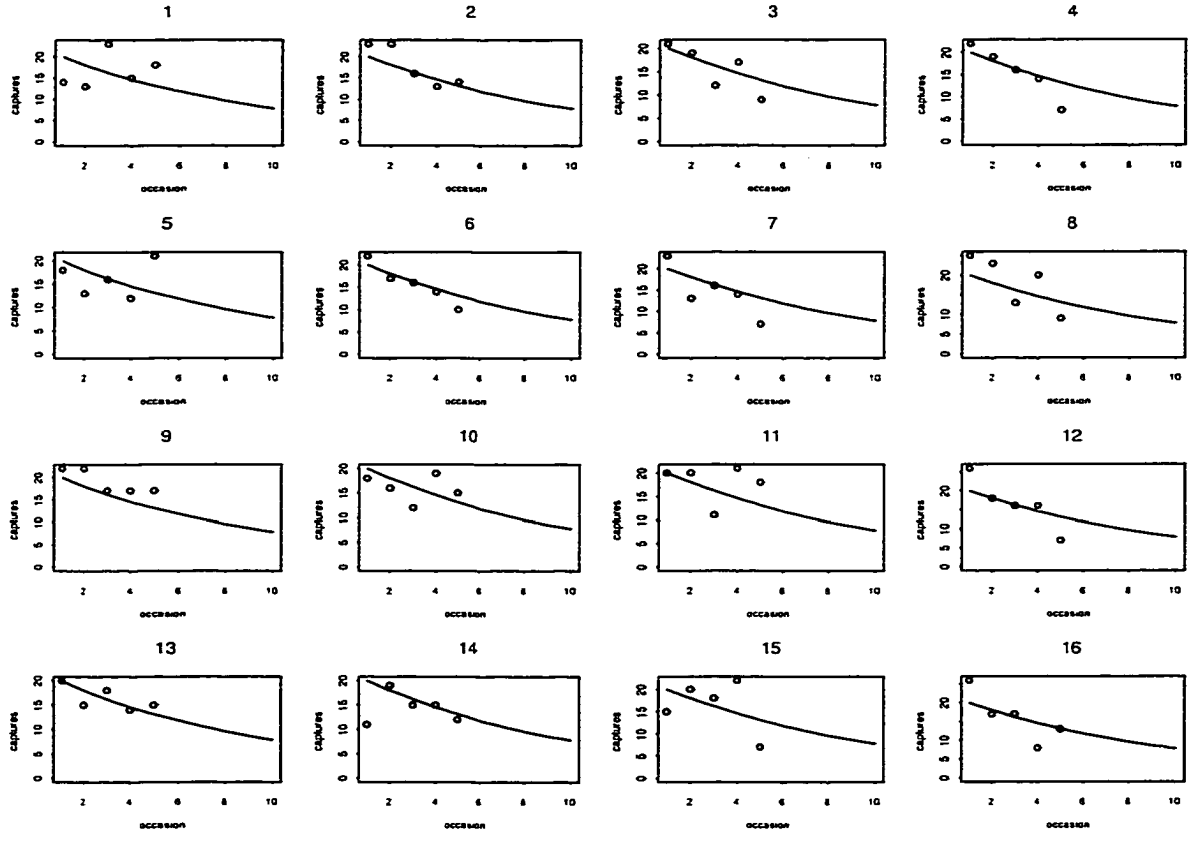


Figure 4.8 Removal data from 16 simulation replicates with $p = 0.1$, $N = 200$, and $k = 5$. Lines represent the expected numbers of first captures with $p = 0.1$. Circles represent the numbers of simulated first captures.

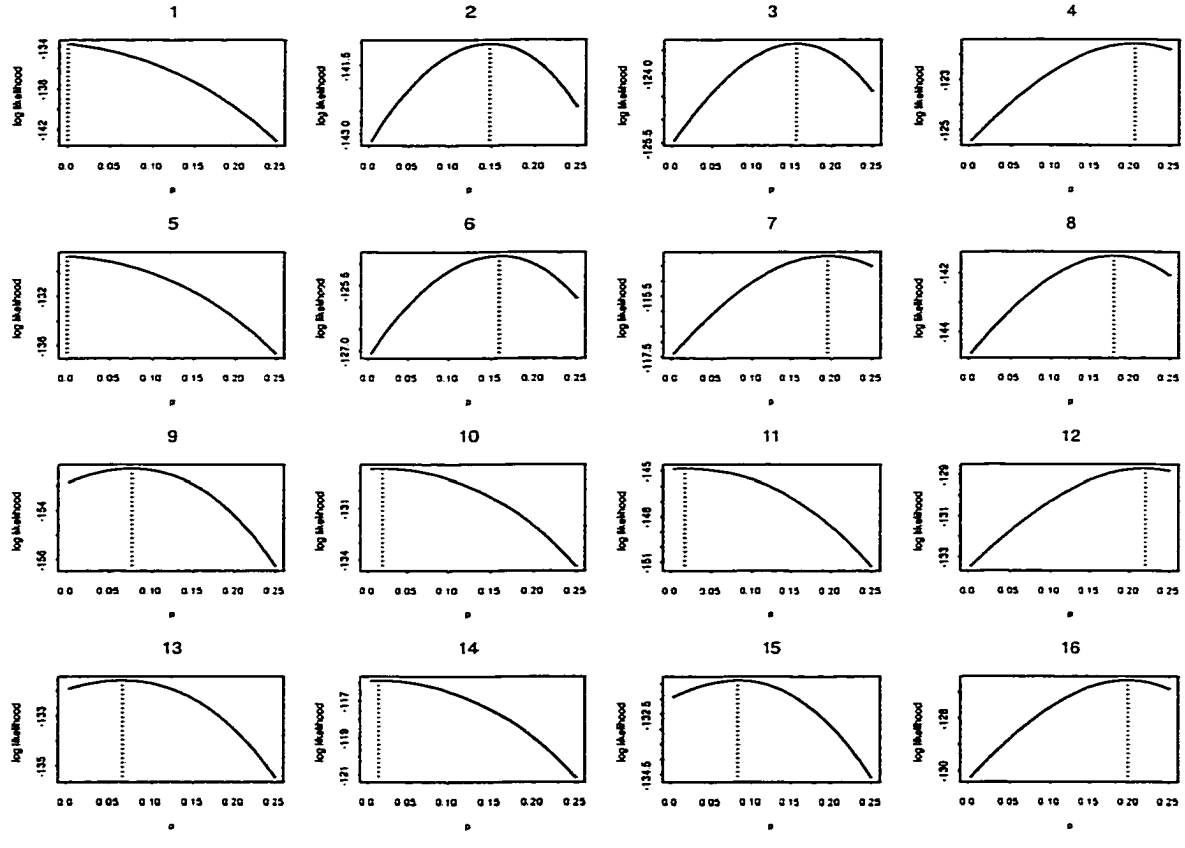


Figure 4.9 Homogeneous geometric removal model log likelihood curves of 16 simulation replicates with $p = 0.1$, $N = 200$, and $k = 5$.

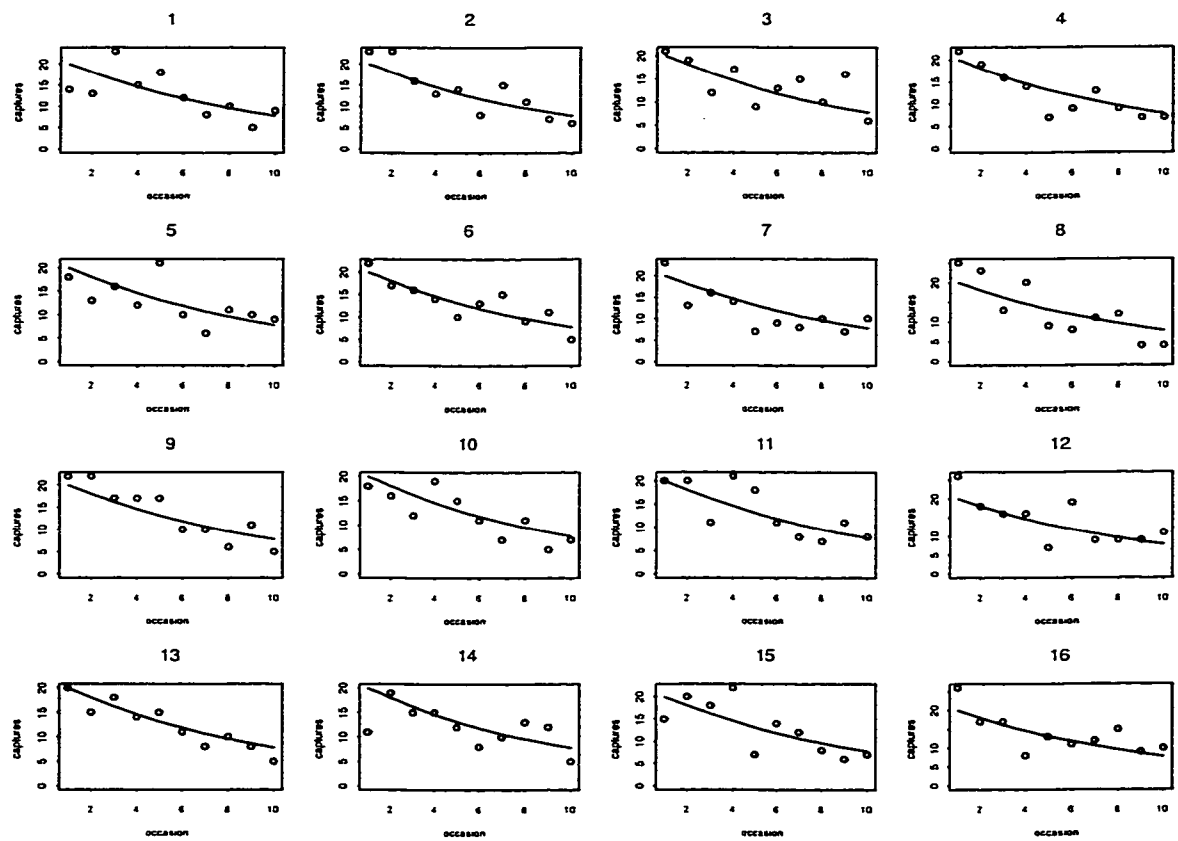


Figure 4.10 Removal data from 16 sumulation replicates with $p = 0.1$, $N = 200$, and $k = 10$. Lines represent the expected numbers of first captures with $p = 0.1$. Circles represent the numbers of simulated first captures.

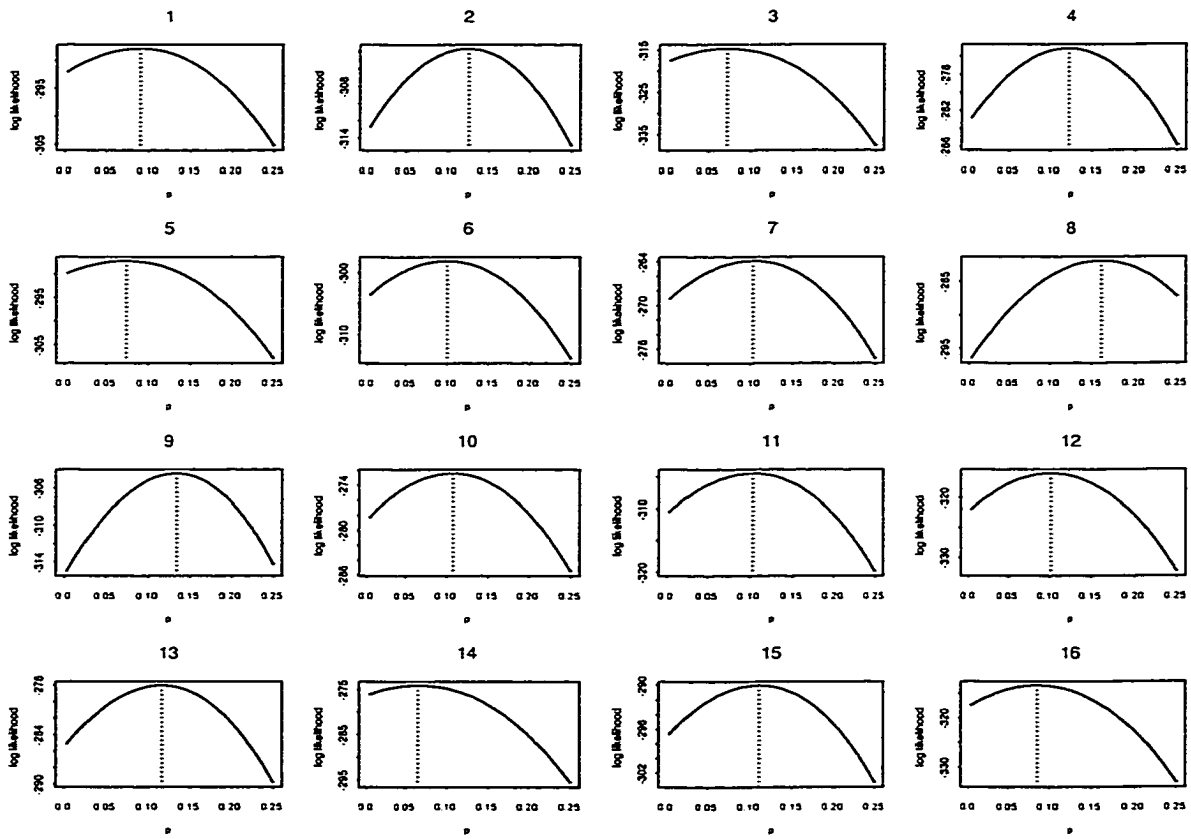


Figure 4.11 Homogeneous geometric removal model log likelihood curves of 16 simulation replicates with $p = 0.1$, $N = 200$, and $k = 10$.

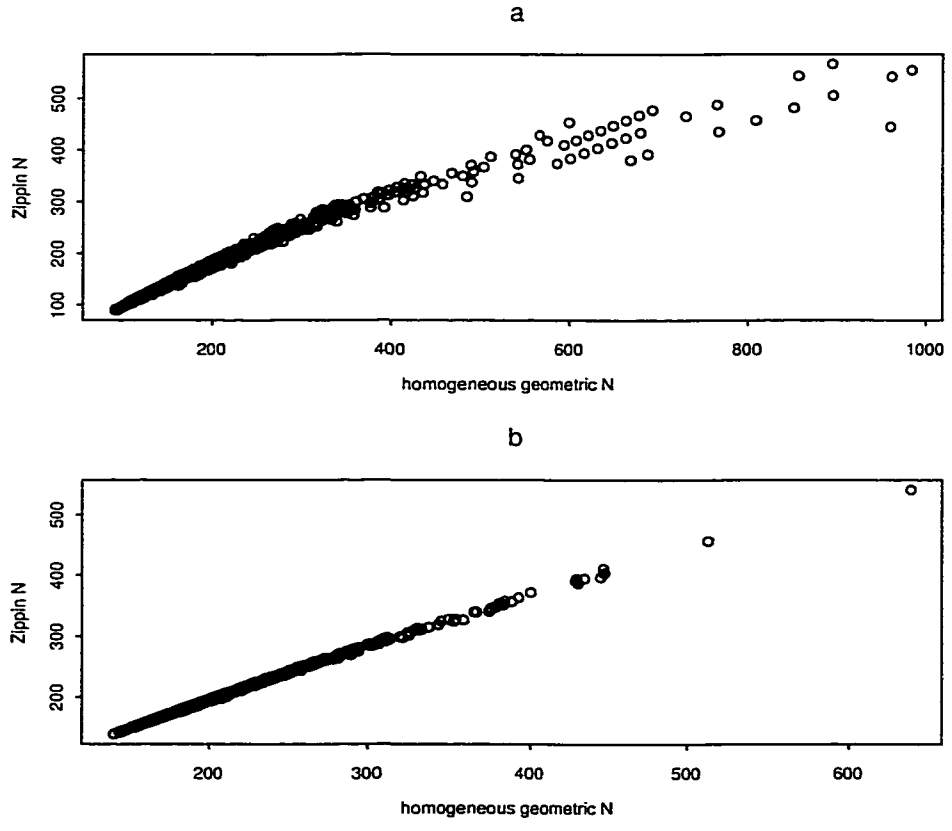


Figure 4.12 Plots of Zippin $\widehat{N}$ versus homogeneous geometric $\widehat{N}$ for simulated data with $N = 200$, $k = 5$, a) $p = 0.1$ and b) $p = 0.2$.

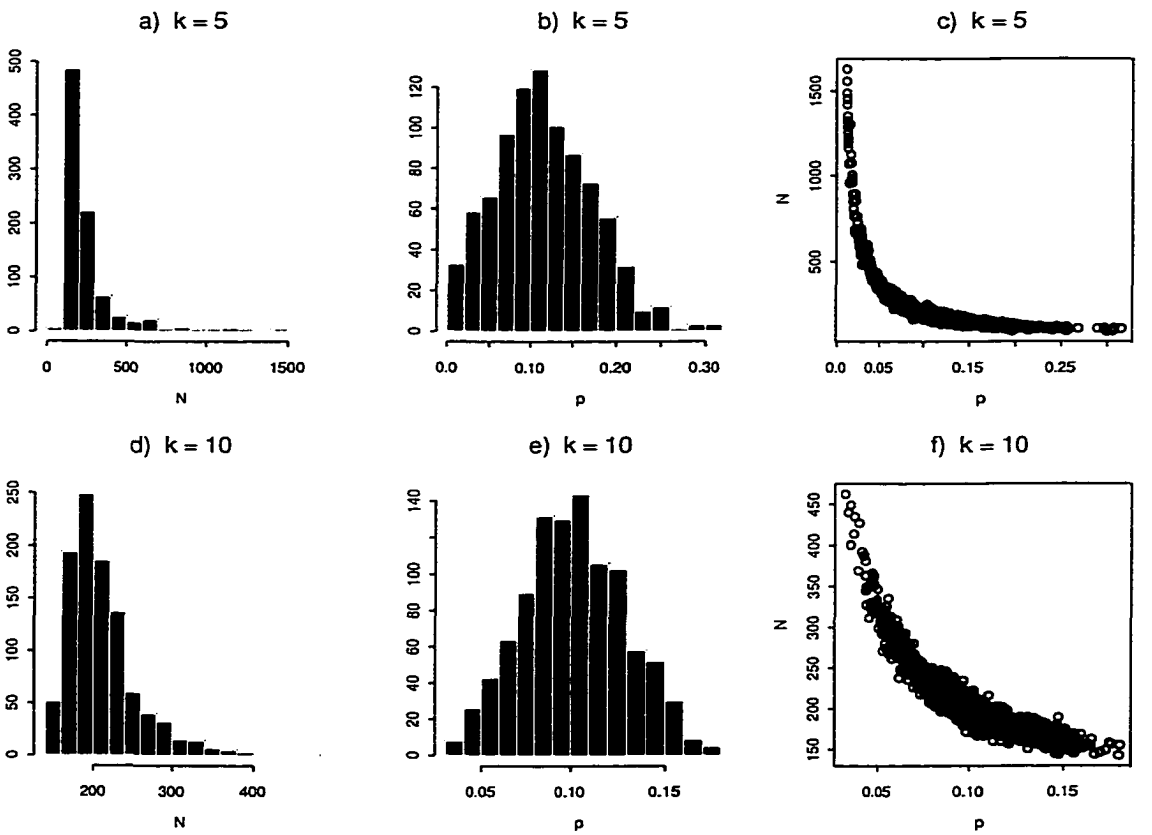


Figure 4.13 Simulations of homogeneous removal with $p = 0.1$ and $N = 200$. Histograms of $\widehat{N}$ and $\widehat{p}$ and plots of $\widehat{N}$ versus $\widehat{p}$ with $k = 5$ are a, b, and c, respectively, and with $k = 10$ are d, e, and f, respectively.

space (Figures 4-13 b and e). These small $\widehat{p}$'s correspond to large $\widehat{N}$'s (Figures 4-13 c and f). When $p = 0.2$ and $N = 200$, $\widehat{N}$ is still skewed (Figures 4-14 a and d) even though $\widehat{p}$ is more symetrically distributed (Figures 4-14 b and d). We also simulated under parameters indicated by the northern pike example in Chapter 1. With $p = 0.355$ and $N = 722$, distributions of $\widehat{N}$ and $\widehat{p}$ are symetric for $k = 5$ and $k = 10$ (Figure 4-15).

Coverage of 95% confidence intervals for model P ranges from 83% to 90% for $p = 0.1$ and 0.2 (Appendix D Tables D1 and D3), though the distribution of $\widehat{N}$ may be skewed. The asymptotic variance for p was calculated as the inverse observed information. The variance of $\widehat{N}$ was calculated using the delta method. For $p = 0.355$ and $N = 722$, the 95% confidence interval coverage was 89%, 67%, and 0% for $k = 5, 10$, and 20, respectively (Appendix A Table A5). At this relatively large level of p, the variance estimator underestimates the true variance. For the pike simulation, for $k = 5, 10$, and 20, median standard errors of $\widehat{N}$ are 14.1, 1.7, and 0.04 which are substantially smaller that the sample RMSE's 17.8, 3.5, and 0.4, respectively (Appendix D Table D5). Other methods for variance estimation need to be explored.

We investigate the effect of small homogeneous probabilities of first capture on the incidence of data that falsely indicate heterogeneity among animals. For 3 simulations of homogeneous populations (Table 4.2), we estimated N from the homogeneous geometric removal model (P), the restricted ($p_2 = 1$) geometric finite mixture model (P2eq1), and the 2-group geometric finite mixture model (P2). For each replicate, the model chosen for estimation had the smallest AIC among models with estimates less than $5N$ and non-significant ($\alpha = 0.05$) goodness of fit (GOF) tests. Model P2eq1 was selected at most 8.4% of the time. Under a homogeneous removal model, the number of first captures can be higher than expected just by chance. This situation may lead to selection of the P2eq1 model. In this case, p_1 represents the capture probability of animals not captured on the first occasion. The P2eq1 model is still representative of homogeneous removal after occasion 1 (i.e., all group 2 animals are removed on occasion 1). Among the 3 simulations, model P2 was not selected when $k = 5$. For $k = 10$, P2 was chosen at most 1% of the time. These small selection rates indicate that a truely homogeneous population is not likely to be incorrectly identified as being heterogeneous.

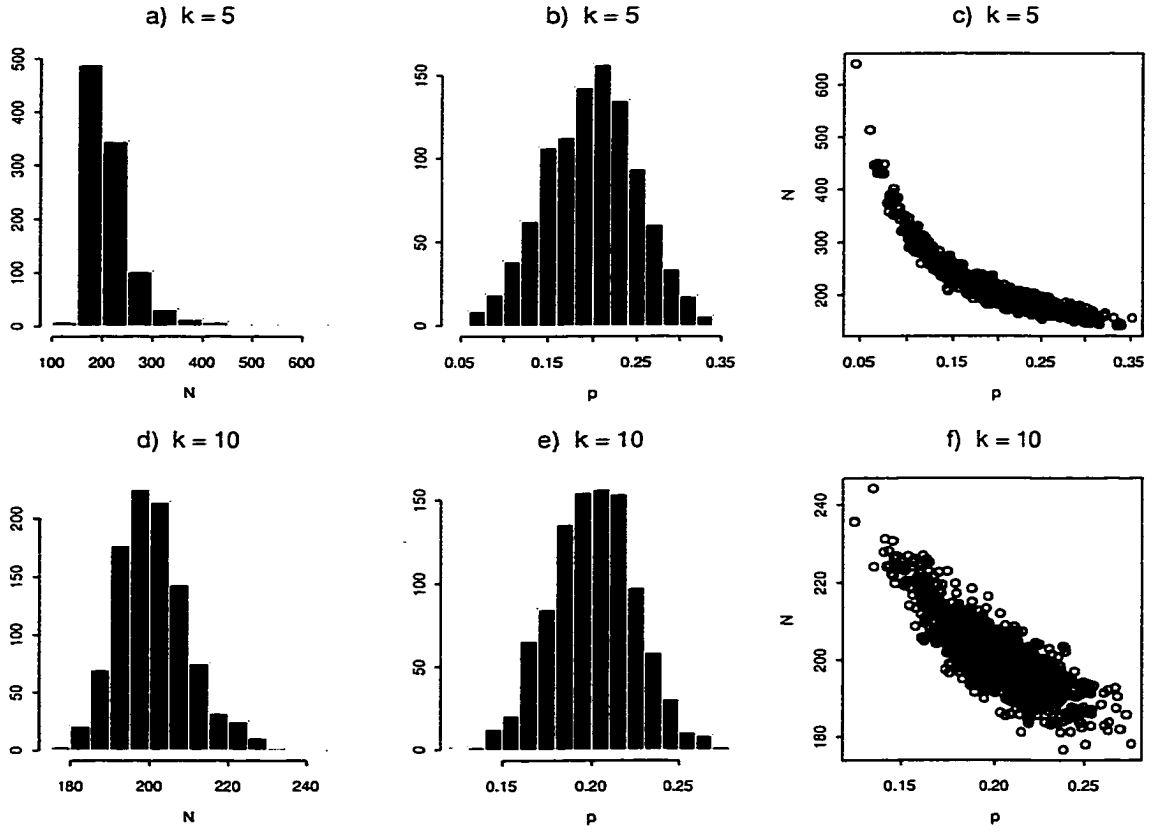


Figure 4.14 Simulations of homogeneous removal with $p = 0.2$ and $N = 200$. Histograms of $\widehat{N}$ and $\widehat{p}$ and plots of $\widehat{N}$ versus $\widehat{p}$ with $k = 5$ are a, b, and c, respectively, and with $k = 10$ are d, e, and f, respectively.

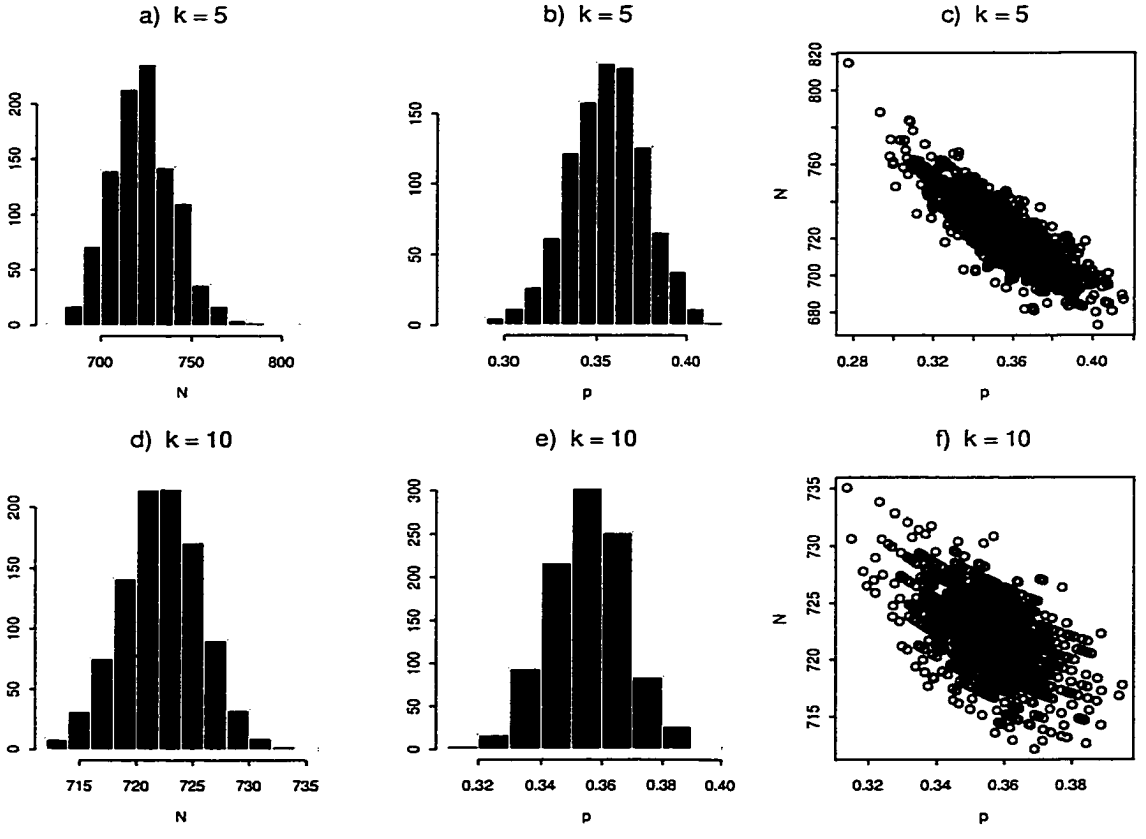


Figure 4.15 Simulations of homogeneous removal with $p = 0.355$ and $N = 722$. Histograms and plots of $\widehat{N}$ and $\widehat{p}$ with $k = 5$ are a, b, and c, respectively, and with $k = 10$ are d, e, and f, respectively.

Table 4.2
AIC *choices among removal models with non-significant GOF and* $\widehat{N} < 5N$.

		p (N)		
		0.355	0.2	0.1
k	method*	(722)	(200)	(200)
5	P	895	878	780
	P2eq1	82	62	12
	P2	0	0	0
	NA	23	60	207
10	P	869	883	888
	P2eq1	84	73	67
	P2	10	3	4
	NA	37	41	41

*methods: P = homogeneous geometric removal, P2eq1 = restricted ($p_2 = 1$) geometric finite mixture, P2 = 2-group geometric finite mixture, NA = no estimate.

4.5 Heterogeneous Population

In the previous section, simulations were limited to populations with 2 groups and no behavioral response to first capture ($\mathbf{p = c}$). Here, we simulate 4 populations to investigate the effects of heterogeneity and behavioral response to first capture on performance of estimators. Each of the four simulations in this section have 1000 replicates. Parameter values and population size are given in Table 4.3. In the last section, one of the simulated homogeneous populations had $p = 0.2$ and Figure 4-14 showed the distribution of estimates of N (Appendix D Table D3). Here, simulation A has a similar population in that the smallest probability of first capture is 0.2. Comparison of distributions of estimates of N between these two simulations may suggest what characteristics of estimators are due to heterogeneity.

The other simulations provide examples of estimation under known levels of heterogeneity and behavioral response to first capture. Simulation B has 3 groups with slight differences in capture probability. The goal of simulation B is to determine if model PeqC3 can differentiate these groups successfully. Parameters for the taxi and vole simulations are like estimates found in Chapter 3. The goal of the Taxi simulation is to observe the effects of slight hetero-

Table 4.3
Parameters for Simulations

name	N	group	π	p	c
A	200	1	0.7	0.2	0.2
		2	0.3	0.9	0.9
B	100	1	0.333	0.1	0.1
		2	0.333	0.2	0.2
		3	0.333	0.5	0.5
Taxi	420	1	0.978	0.104	0.132
		2	0.022	1	0.267
Vole	133	1	0.704	0.117	0.493
		2	0.158	0.883	0.775
		3	0.138	1	0.223

geneity on the distributions of estimators. The vole simulation has strong heterogeneity with varied behavioral response to first capture. Previously for this combination of effects, removal models were used for simulation and estimation (Otis et.al. 1978), and recaptures were not simulated. Our models PC2 and PC3 use recapture data for estimation of N in the presence of heterogeneity and behavioral response to first capture. The purpose of the vole simulation with recaptures is to determine if these recapture estimators are more efficient than removal estimators.

4.5.1 Simulation A

In this simulation, the population of $N = 200$ has two groups, 70% have $p_1 = c_1 = 0.2$ and 30% have $p_2 = c_2 = 0.9$. These animals have no behavioral response to first capture. To examine the effects of heterogeneity on performance of estimators, we contrast results of simulation A to the homogeneous simulation from section 4.4 with $N = 200$ and $p = c = 0.2$. The median numbers of animals caught in 5 occasions are 135 for the homogeneous simulation and 154 for the heterogeneous simulation. Comparison of boxplots for the heterogeneous simulation in Figure 4-16 to the corresponding boxplots for the homogeneous simulation in Figure 4-17 shows that distributions of $\widehat{N}$ for the homogeneous removal estimator (p), Zippin (Z), and the multinomial mark-recapture null model estimator (null) are much tighter in the heterogeneous

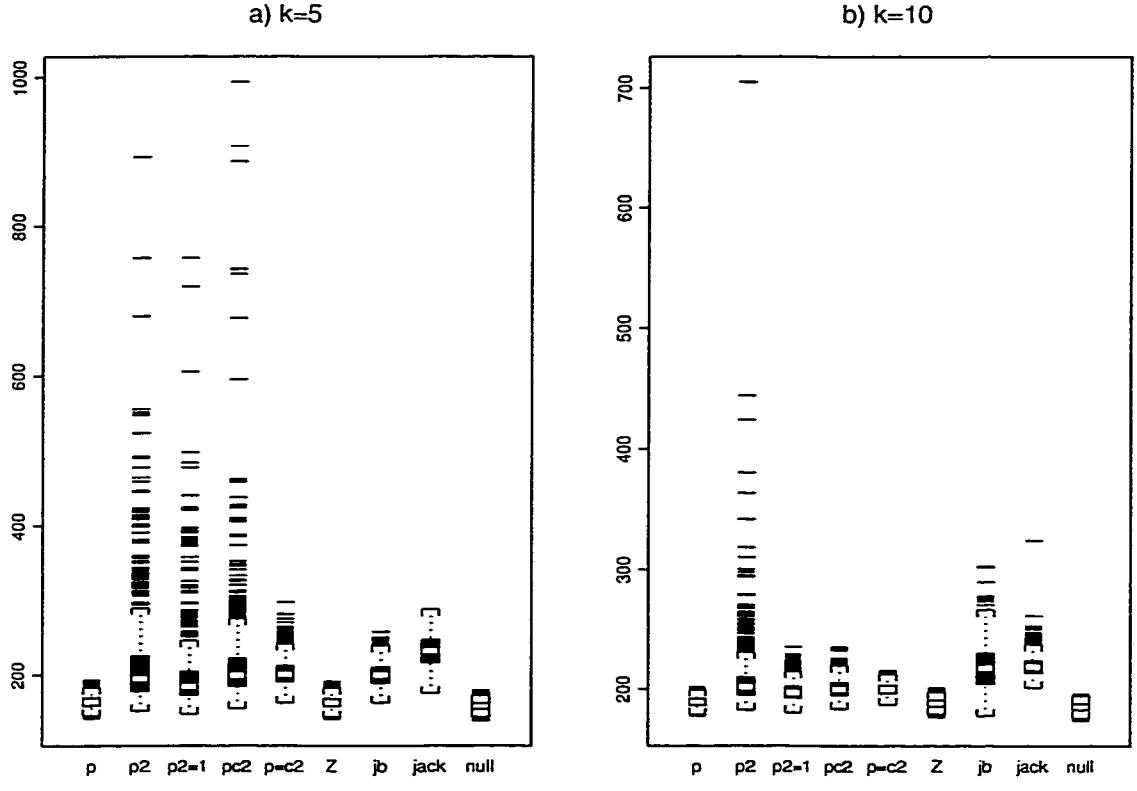


Figure 4.16 Simulation A, N = 200, upper limit = 1000; p: homogeneous geometric removal; p2: 2-group geometric removal; p2=1: 2-group geometric ($p_2 = 1$) removal; pc2: 2-group mark-recapture (MR) mixture; p=c2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; Z: Zippin (1956); jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

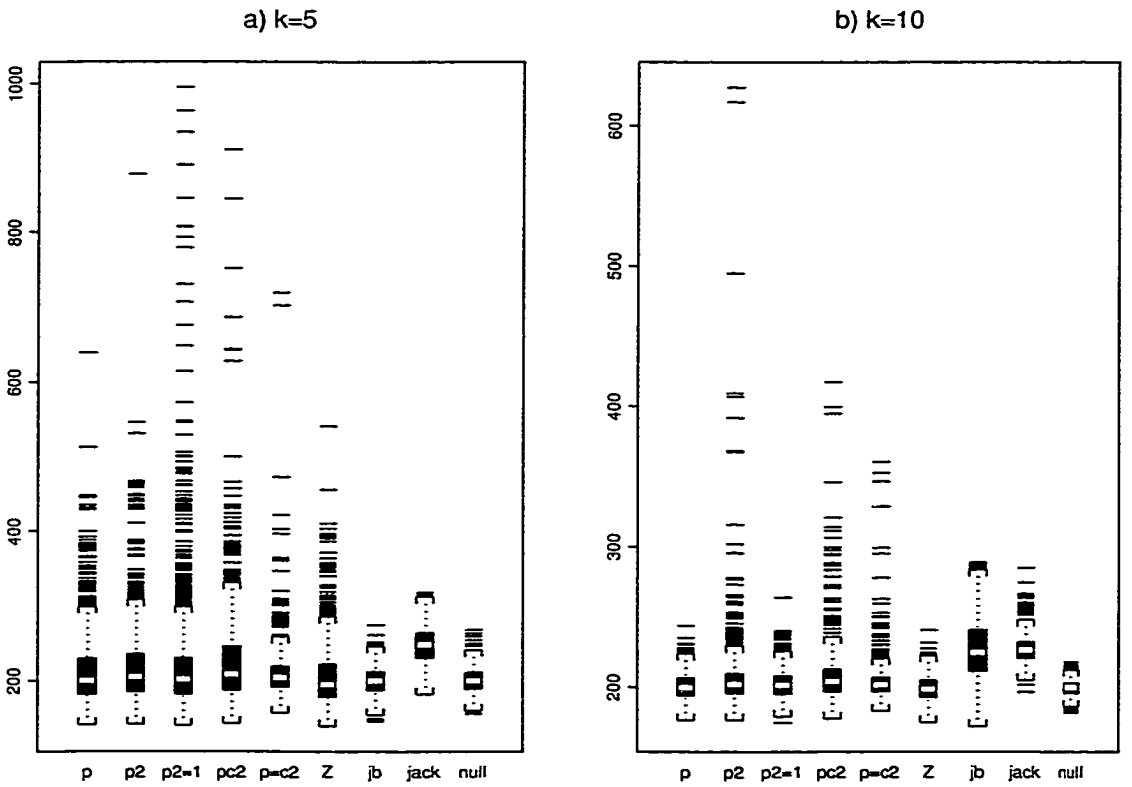


Figure 4.17 Simulation of a homogeneous population, $p = 0.2$, $N = 200$, upper limit =1000; p: homogeneous geometric removal; p2: 2-group geometric removal; p2=1: 2-group geometric ($p_2 = 1$) removal; pc2: 2-group mark-recapture (MR) mixture; p=c2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; Z: Zippin (1956); jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

simulation (Figure 4-16), but have negative bias (Appendix E Tables E2 and E3).

For finite mixture models which allow heterogeneity, PC2 and PeqC2 estimates have smaller RMSE's in the heterogenous simulation (Appendix E Table E3) than in the homogeneous simulation (Appendix D Table D4). This relates to the difficulty of estimating extra parameters for heterogeneity when the underlying model is homogeneous. The likelihood surface is very flat in directions of parameters that are not part of the underlying model. In Figure 4-16a, the finite mixture estimators (p2, p2=1, pc2, and p=c2) show some skewness (with outliers and non-symmetric boxplots). This is reduced when $k = 10$, Figure 4-16b. Large estimates of N are related to estimates of the smallest probability of first capture close to 0. With more data from occasions 6 through 10, estimates of the smallest probability of first capture are farther from 0 and skewness in $\widehat{N}$ is reduced.

All estimators in Figure 4-18 have smaller RMSE's at $k = 10$ than at $k = 5$, except the Pollock and Otto (1983) jackknife (jb). At $k = 5$, model PeqC2 and jb estimators are most efficient. At $k = 10$, model PeqC2 is most efficient. In this simulation, the restriction $\mathbf{p} = \mathbf{c}$ helps with estimation of N, especially for $k = 5$. The model PeqC2 estimator is more efficient than the model PC2 estimator.

4.5.2 Simulation B

This simulation has 1000 replicates on a population with 3 groups, each having 1/3 of the animals and no behavioral response to first capture (i.e., $\mathbf{p} = \mathbf{c}$). The probabilities of first capture (p) are 0.1, 0.2, and 0.5 (Table 4.3). The goal of the simulation was to determine if model PeqC3 successfully differentiates these groups. We successfully estimated parameters of the PeqC3 model for 543 replicates with $k = 5$ and 739 replicates with $k = 10$. The true difference between $p_2 = 0.2$ and $p_1 = 0.1$ is 0.1, though at $k = 5$, the third quartile for the difference $\widehat{p}_2 - \widehat{p}_1$ is 0.014, well below the true value. For conditions of this simulation, the PeqC3 model does not successfully differentiate between groups with capture probabilies of 0.1 and 0.2. At $k = 10$, the estimated difference has median < 0.001 and third quartile $= 0.17$. For more than half of the replicates with estimates, the 3-group model does not provide more

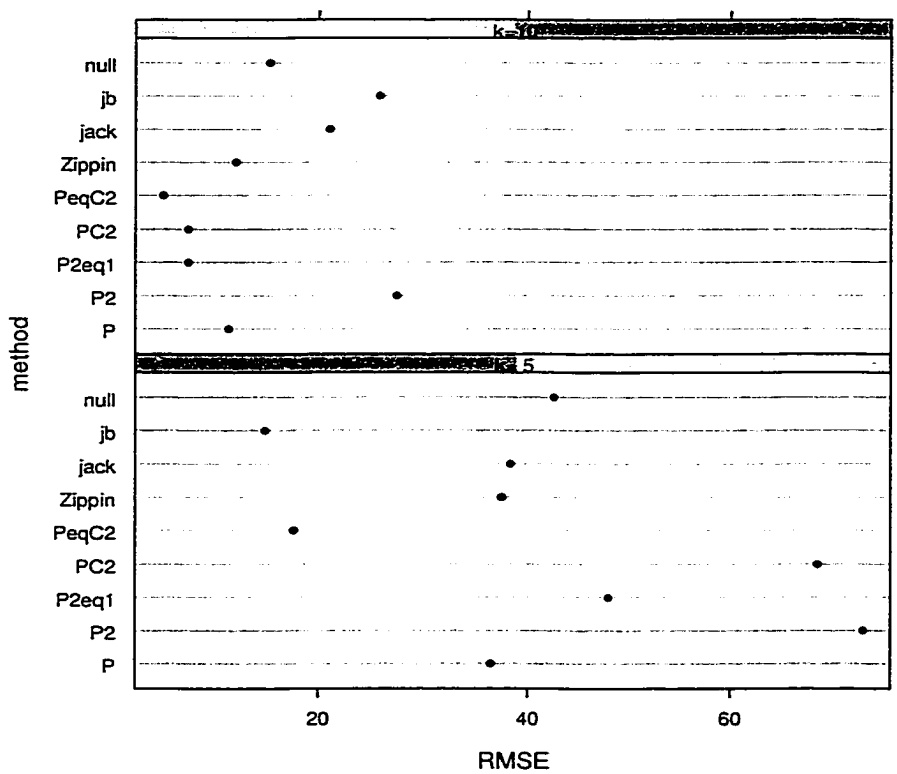


Figure 4.18 Simulation A RMSE's; null: Otis et. al. (1978) multinomial; jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; Zippin: Zippin (1956); PeqC2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; PC2: 2-group mark-recapture (MR) mixture; P2eq1: 2-group geometric ($p_2 = 1$) removal; P2: 2-group geometric removal; P: homogeneous geometric removal.

information than the 2-group model PeqC2. For 1- and 2-group models, the skewness among our estimators at $k = 5$ (Figure 4-19) is not reduced by using 5 additional occasions ($k = 10$, Figure 4-19b). At $k = 10$, the Burnham and Overton (1978) jackknife estimator (jack) has more positive outliers than at $k = 5$, which is contrary to what one would expect with an increase in data.

All estimators shown in Figure 4-20 (Appendix E Tables E4 and E5) have smaller RMSE's at $k = 10$. At $k = 5$, the finite mixture estimators are less efficient (larger RMSE's) than the homogeneous geometric removal estimator (P) or the Zippin (1956) estimator. At $k = 10$, more data are available and the fit of model PeqC2 is much improved, but the model P removal estimator is more efficient.

In Figure 4-20 at $k = 5$, model PeqC2 is the least efficient estimator, though at $k = 10$ it is improved and is more efficient than jb or jack. At $k = 10$, the estimators of models P, P2eq1, Zippin, and null are more efficient than estimators that allow more heterogeneity. This suggests that the effects of heterogeneity for this simulation can be overcome by a long series of sampling occasions. As more occasions are sampled, more animals with smaller probabilities of first capture contribute to the data and offset the influence of animals caught on early occasions with high probabilities of first capture. At $k = 10$, 85% of the population is expected to be caught. When p is similar among groups as here, modeling the mixture to characterize the uncaptured 15% of the population is of little advantage for estimating N. Perhaps values of p among the groups are too similar to split out by fitting models to the data.

4.5.3 Taxi

In the taxi cab example of Chapter 2, the PC2 model has $\widehat{N} = N = 420$, though diagnostics indicate that a homogeneous model is better. Heterogeneity of the taxi model with $p_2 = 1$ is expressed in first captures at the first occasion. Probabilities of first capture among cabs not seen on the first occasion are homogeneous ($p_1 = 0.104$). What are characteristics of estimators on this population with slight heterogeneity? Estimators of geometric removal models show strong skewness when $k = 5$ (Figure 4-21a), like the distributions of estimates for simulation

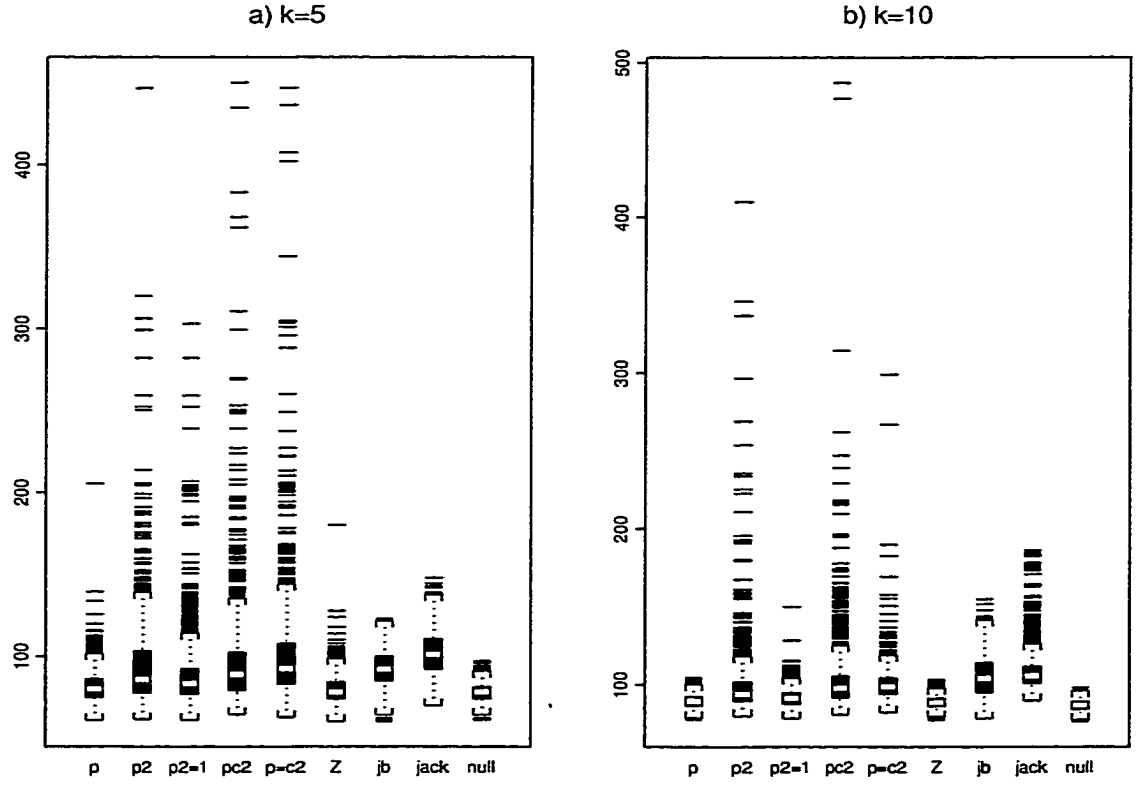


Figure 4.19 Simulation B, N=100, upper limit = 500; p: homogeneous geometric removal; p2: 2-group geometric removal; p2=1: 2-group geometric ($p_2 = 1$) removal; pc2: 2-group mark-recapture (MR) mixture; p=c2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; Z: Zippin (1956); jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

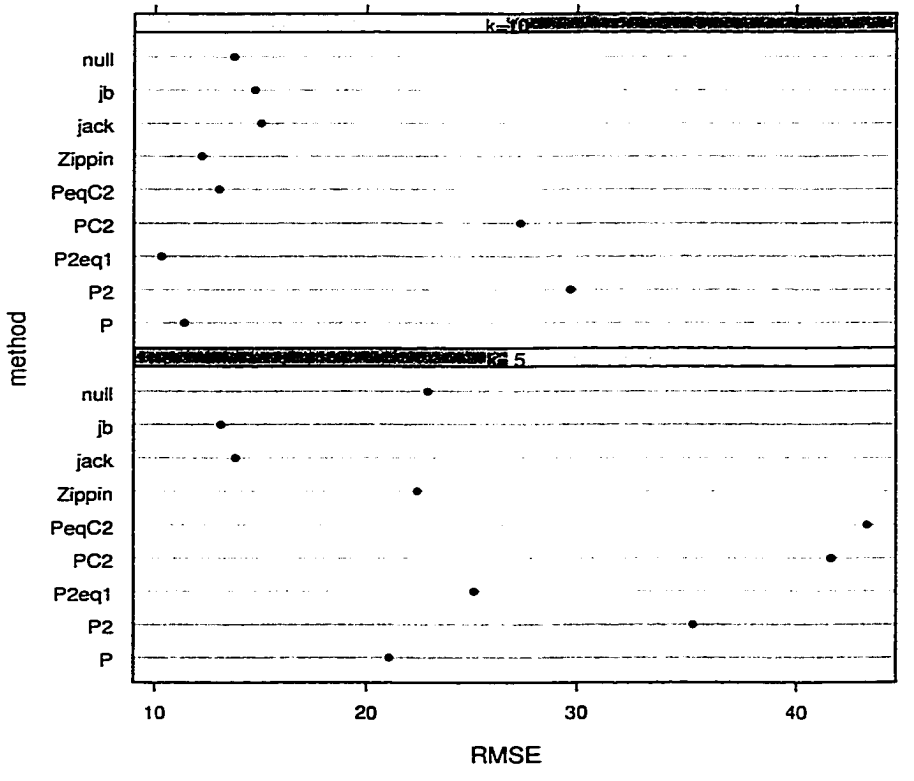


Figure 4.20 Simulation B RMSE's; null: Otis et. al. (1978) multinomial; jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; Zippin: Zippin (1956); PeqC2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; PC2: 2-group mark-recapture (MR) mixture; P2eq1: 2-group geometric ($p_2 = 1$) removal; P2: 2-group geometric removal; P: homogeneous geometric removal.

B. For models P and P2eq1, this skewness is reduced by the addition of 5 occasions ($k = 10$, Figure 4-21b). Since the underlying heterogeneity is slight, this pattern relates more to the small probability of first capture than to heterogeneity. In Figure 4-22, (Appendix E Tables E6 and E7) RMSE's of jack are relatively small, but at $k = 10$ it is larger than at $k = 5$ (i.e., the estimator is not better with the 5 additional occasions). At $k = 10$, homogeneous model estimators perform nearly as well (small RMSE's) as the jackknife estimators (jb, jack).

4.5.4 Vole

Unlike the previous 2 simulations, this simulation is on a 3 group population of 133 voles (Chapter 3) with strong heterogeneity in probabilities of first capture and heterogeneity in behavioral response to first capture (Table 3.3). The geometric finite mixture model estimators have more positive outliers than the jackknife estimators, Figure 4-23. At $k = 5$, the model PC3 estimator has estimates $< 3N$ for only 44% of the 1000 replicates. This model has 8 parameters, so the difficulty with estimation from 5 occasions is not surprising. Model PC2 has more negative bias, but its RMSE is smaller than that of PC3, at $k = 5$ and 10 (Figure 4-24). For both $k = 5$ and 10, the PC2 estimator is more efficient than the P2 estimator. This shows that estimation of $\mathbf{c}$ improves estimation of N, even though $\mathbf{p} \neq \mathbf{c}$. At $k = 10$, the PC2 estimator is more efficient than the jackknife estimators (Figure 4-24, Appendix E Tables B8 and B9). The RMSE of model PC3 is the largest (Figure 4-24), however with careful use of plots of fit (as shown in Chapter 3), tests, and indices for model selection, this model can be used effectively. At $k = 5$, the Pollock and Otto jackknife removal (jb) estimator had a root mean squared error (RMSE) of 27.0, while the 2-group mark-recapture finite mixture (PC2) estimator had RMSE = 44.6. However,at $k = 10$, the relative efficiency of our estimator was better, RMSE(Jb) = 18.1 and RMSE(PC2) = 17.9.

4.5.5 Model Selection

In Chapters 2 and 3, we present examples with figures that display the fit and provide an impression of how the models are working. Here, we investigate a "turnkey" approach to

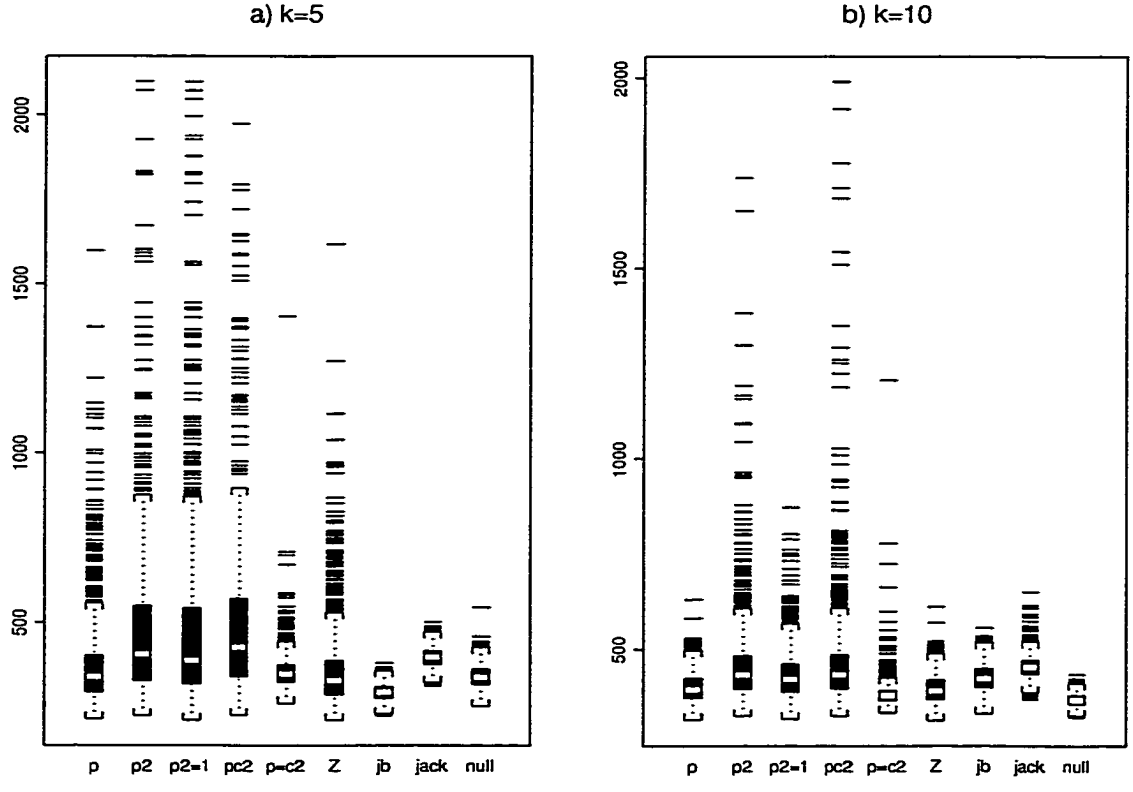


Figure 4.21 Simulation Taxi, N = 420, upper limit = 2100; p: homogeneous geometric removal; p2: 2-group geometric removal; p2=1: 2-group geometric ($p_2 = 1$) removal; pc2: 2-group mark-recapture (MR) mixture; p=c2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; Z: Zippin (1956); jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

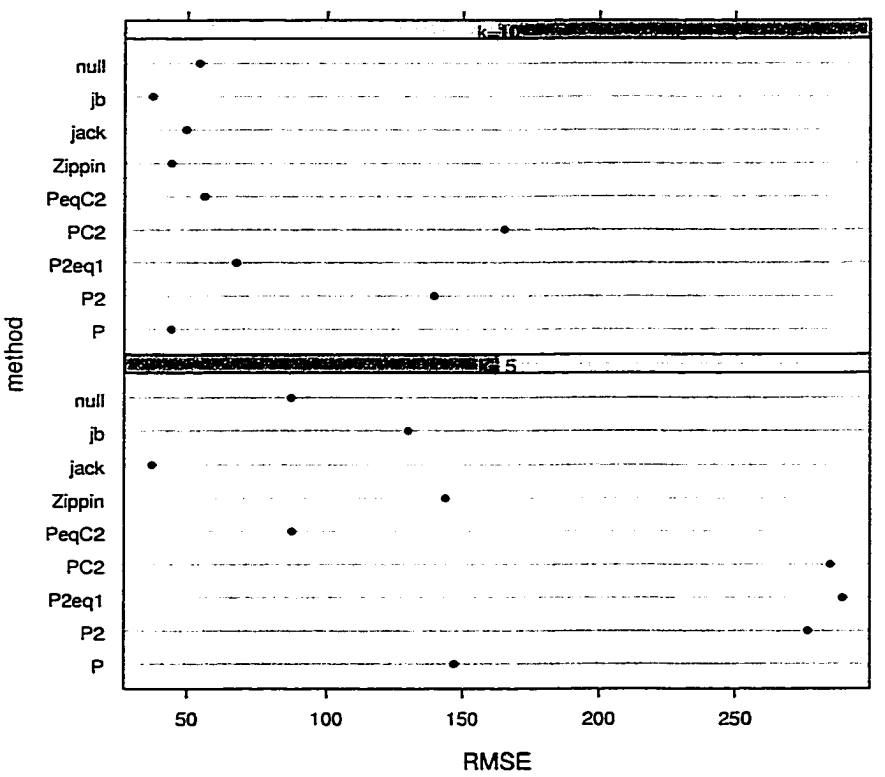


Figure 4.22 Simulation Taxi RMSE's; null: Otis et. al. (1978) multinomial; jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; Zippin: Zippin (1956); PeqC2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; PC2: 2-group mark-recapture (MR) mixture; P2eq1: 2-group geometric ($p_2 = 1$) removal; P2: 2-group geometric removal; P: homogeneous geometric removal.

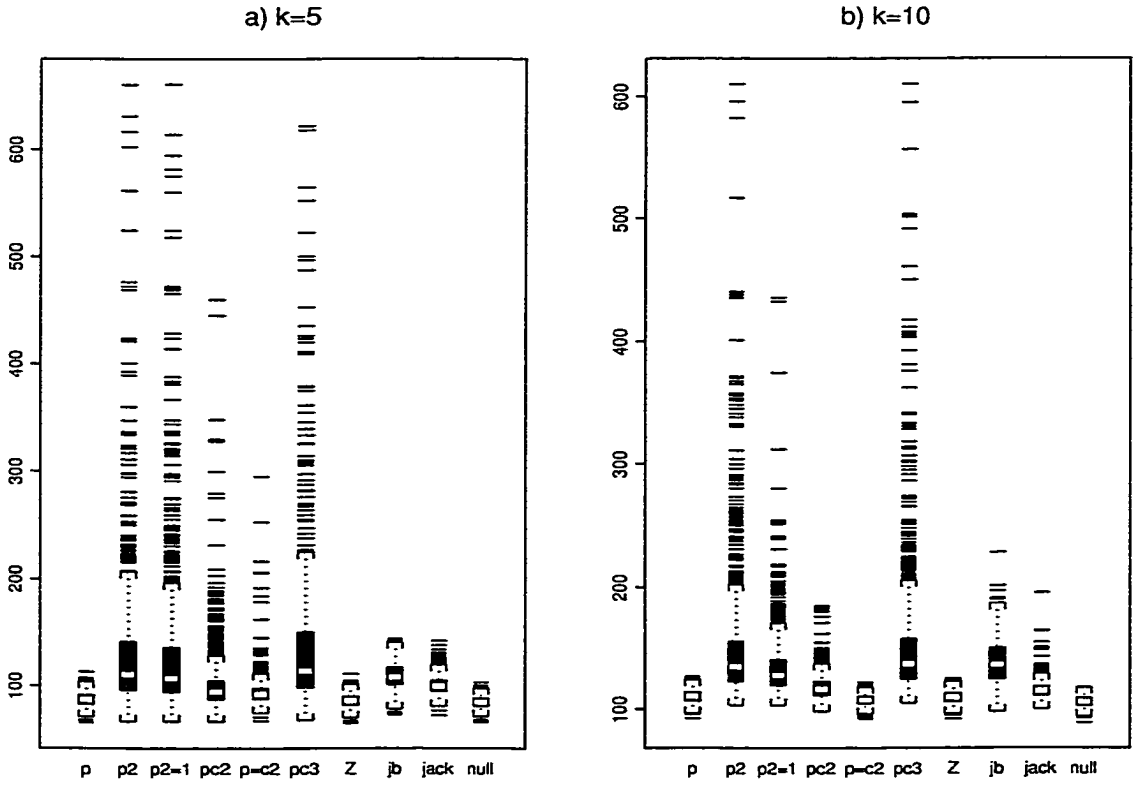


Figure 4.23 Simulation Vole, N=133, upper limit = 665; p: homogeneous geometric removal; p2: 2-group geometric removal; p2=1: 2-group geometric ($p_2 = 1$) removal; pc2: 2-group mark-recapture (MR) mixture; p=c2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; Z: Zippin (1956); jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

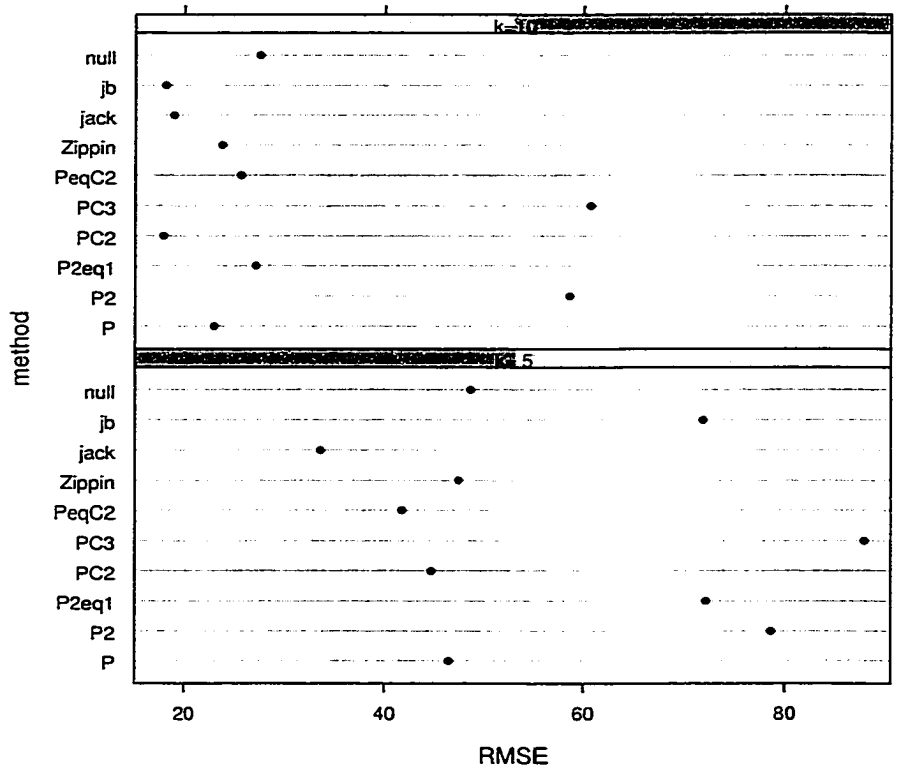


Figure 4.24 Simulation Vole RMSE's; null: Otis et. al. (1978) multinomial; jb: Pollock and Otto (1983) jackknife; jack: Burnham and Overton (1978) jackknife; Zippin: Zippin (1956); PeqC2: 2-group MR ($\mathbf{p} = \mathbf{c}$) mixture; PC2: 2-group mark-recapture (MR) mixture; P2eq1: 2-group geometric ($p_2 = 1$) removal; P2: 2-group geometric removal; P: homogeneous geometric removal.

Table 4.4.
Number of AIC *choices among models with non-significant* GOF *and* $\widehat{N} < 5N$.

		removal			mark-recapture			
simulation	k	P	P2	P2eq1	PC	PC2	PeqC2	PC3
A	5	49	0	928	0	137	824	
	10	0	216	762	0	146	799	
B	5	829	0	127	388	118	398	
	10	600	37	310	0	170	753	
taxi	5	743	0	160	798	52	84	
	10	627	8	327	530	290	141	
vole	5	38	0	806	6	261	329	157
	10	0	90	880	0	8	0	866

model selection that does not incorporate the investigators knowledge nor interpretation of plots. The algorithm selects by AIC among geometric removal models (P, P2, P2eq1) with $\widehat{N} < 5N$ and non-significant GOF tests ($\alpha = 0.05$). We also looked at selection among mark-recapture models (PC, PC2, PeqC2, PC3), Table 4.4. At $k = 5$, none of the simulations had replicates with model P2 as the choice. This model with 3 parameters is difficult to fit with frequencies of first capture at only 5 occasions. Model PC3 was not used in simulations A, B, and Taxi, because MLE's of PC3 were difficult to find.

In simulation A, data were created with parameters $\mathbf{p}' = \mathbf{c}' = (0.2, 0.9)$ and $\pi' = (0.7, 0.3)$, Table 4.4. Among geometric removal models (P, P2, P2eq1), P2eq1 is chosen most often because p_2 is close to 1. When the number of occasions was increased from 5 to 10, P2 was chosen in 22% of the replicates. The increase in data enables estimation of p_2 at levels less than 1. Among the mark-recapture models, the restriction $\mathbf{p} = \mathbf{c}$ in the 2-group model was chosen in 84% of the replicates. Here, conditions indicated by model selection match the underlying model.

Parameters of simulation B are $\mathbf{p}' = \mathbf{c}' = (0.1, 0.2, 0.5)$ and $\pi' = (0.33, 0.33, 0.33)$. Among the removal models, P is chosen most often. Among the mark-recapture models at $k = 5$, the choice of 43% of the replicates indicate homogeneity (PC). Elements of $\mathbf{p}$ are too close to be differentiated without 10 occasions and recapture data. The condition $\mathbf{p} = \mathbf{c}$ was correctly

identified 82% of the time when $k = 10$. Though model selection often indicates homogeneity, the biases of model P, Zippin, and null model estimators at $k = 5$ are -18.7, -20.7, and -22.2, respectively, for a population of 100 (Appendix E Tables E4 and E5).

The taxi simulation is created from 2 groups, but the heterogeneity is slight because the largest group comprised 97.8% of the population. The underlying model has slight behavioral response to first capture ($\mathbf{p} \neq \mathbf{c}$). Homogeneous models P and PC were chosen most often, Table 4.4. When 10 occasions were used, heterogeneous models (P2eq1, PC2) were selected more often than when 5 occasions were used. More data allow a greater ability to discriminate between models.

The vole simulation has heterogeneity in behavioral response to first capture and in probabilities of first capture. Among the removal models, P2eq1 is chosen most often, because 29% of the population is in a group with $p > 0.88$ which is close to 1. Among the mark-recapture models at $k = 10$, PC3 was chosen in 99% of the replicates. According to the underlying model, 80% of the animals are expected to be caught at least once in the 10 occasions. This large sampling effort is the price for sufficient data to model heterogeneity among 3 groups. In general, small capture probabilities, short surveys ($k = 5$), or small differences in parameters among groups lead to data for which differentiation of groups can be difficult.

4.6 Discussion

Our estimators do not always produce estimates. In the search for MLE's we set a limit of 500 to the number of steps that the algorithm would make before finding the MLE or abandoning the search. Also, for simulations, we dismissed use of $\widehat{N}$'s greater than $5N$. The 3 parameters of model P2 are sometimes difficult to estimate from removal data on 5 occasions. Simulations of homogeneous populations illustrate that when low proportions of populations are observed during surveys (due to small k or p), patterns of data can occur that make estimation difficult though model assumptions are not violated. When the model used for estmation has more groups than the population, then estimation is poor. Models PC3 and PC2 had poor estimation on data from populations with 2 groups and 1 group, respectively. One

would expect a likelihood surface to be relatively flat across dimensions of parameters that are not inherently part of the population and capture process. By contrast, the jackknife estimators have closed forms and have estimates for every replicate in every simulation presented here, though they do not describe the capture process as our models do.

For simulations having the smallest p equal to 0.1, distributions of estimates for N were skewed, especially when $k = 5$ (Figure 4-13). Also, the comparison of the Zippin estimator and the homogeneous geometric model (P) estimator in Figure 4-12 shows that when the underlying model is homogeneous, the model P estimate for a given data set is larger than the Zippin estimate. Perhaps development of a Zippin finite mixture would provide estimators with less positive bias when capture probabilities are heterogeneous.

Among simulations at $\pi = 0.5, 0.7, 0.9$, $p_1 = 0.1, 0.2$, $p_2 = 0.5, 0.9$, and $k = 5, 10$, the model PeqC2 estimator is more efficient than the Burham and Overton (1978) jackknife estimator when $k = 10$. Under conditions of heterogeneity, increases in sampling may not produce Burnham and Overton (1978) jackknife estimates of N closer to the true N, but the increase does improve maximum likelihood estimates. Also, comparison of RMSE's suggest that for removal data with heterogeneous probability of first capture among animals, the Pollock and Otto (1983) jackknife estimator is better than our removal estimators. However, if one plans to use this jackknife removal estimator, increasing the number of occasions may not improve estimates, because its RMSE's do not monotonically decrease as occasions increase (Figure 4-3).

For our estimators, differentiation of groups through estimation and model selection is more difficult when groups have similar capture probabilities or when the a small portion of the population is observed. The model choice by GOF tests, AIC, and upper limits on $\widehat{N}$ does not mislead to indicate a level of heterogeneity that is greater that of the underlying model.

Natarajan and McCulloch (1999) developed a test for heterogeneity in nest survival data. Nest survival is time to event data and so is the time to first capture in removal data. This test could be adapted to test for heterogeneity in time of first capture among animals captured at least once in k occasions.

CHAPTER 5. SUMMARY

For data from a Schnabel census with k occasions, we modeled animals as belonging to discrete groups in which animals have the same probabilities of first capture and recapture. For removal data, Seber (1982) describes a truncated geometric distribution for homogeneous probabilities of first capture. We developed a similar approach with a 2-group geometric finite mixture model for heterogeneous probabilities of first capture. We also developed a version of this model that restricts the probabiltiy of first capture in one group to equal one.

For mark-recapture data, binomial finite mixtures for recaptures were combined with a geometric finite mixture for first captures to better estimate mixture proportions. This model requires the assumption that each animal's group affiliation was not changed by first capture. Among animals that were first caught on occasion x,recaptures were modeled as a binomial finite mixture of v groups with $\tau_u(x)$ as the proportion of animals in group u,

$$\tau_u(x) = \frac{\pi_u p_u (1-p_u)^{x-1}}{\sum_{l=1}^{v} \pi_l p_l (1-p_l)^{x-1}}, \text{ for } u = 1, \ldots, v \text{ and } x = 1, \ldots, k.$$

For one data set, the model potentially has k binomial finite mixtures. Through $\tau_u(x)$, modeling recapture data helps to characterize the mixture and improve estimation of N. Assumptions for this model are comparable to those of M_{bh} in Otis, et al. 1978 (heterogeneity and behavioral response to first capture). For a model that has assumptions comparable to M_h (i.e., no behavioral response to first capture), we fit the model above with the restriction that within each group, probabilities of first capture are equal to probabilities of recapture.

On Carother's (1973) taxi cab data, estimation with a 2-group mark-recapture finite mixture provided a population size estimate, $N = 420$, that exactly matches the number of registered cabs. On meadow vole data, estimation with a 3-group model showed heterogeneity in behavioral response to first capture. In one population, some voles were trap "happy" and

others were trap "shy".

Simulations show that our 2-group mark-recapture finite mixture estimator with restriction for no behavioral response to first capture is more efficient than Burnham and Overton's (1978) jackknife estimator when the smallest probability of first capture is 0.1 and the number of sampling occasions is 10. In a simulation of the vole population above at $k = 5$, the Pollock and Otto jackknife removal (Jb) estimator had a root mean squared error (RMSE) of 27.0, while the full 2-group mark-recapture finite mixture (PC2) estimator had RMSE = 44.6. However,at $k = 10$, the relative efficiency of our estimator was better, RMSE(Jb) = 18.1 and RMSE(PC2) = 17.9. Also, our PC2 model has parameters that help the researcher understand the capture process and the population.

APPENDIX A. EM ALGORITHM FOR REMOVAL ESTIMATION

For EM, a method that allows support of X_i to be $\{1, 2, ...\}$, regard the unobservable X_i's as missing data. Dempster, Laird and Rubin (1977) "regard the unknown number of missing sample points along with their values as being the complete data."

Homogeneous Geometric

The EM algorithm can be used to find the MLE's. Let U be the random variable for the unknown number of missing value points. As in Section 2, $\mathbf{X}$ is a $1 \times N$ vector of independent geometric random variables for times to first capture with no truncation such that X_i has support $\{1, 2, \ldots\}$ for $i = 1, 2, ..., N$ animals. Map $\mathbf{X}$ into observable $\mathbf{Y}$ with support $\{1, 2, \ldots, k\}$ and into unobservable $\mathbf{W}$ with support $\{k+1, k+2, \ldots\}$ such that $y_i = x_i$ for animals $i = 1, 2, ..., m$ and $w_i = x_i$ for animals $i = m+1, m+2, ..., N$. Now, the complete data likelihood from mass function (2.1) can be written as

$$L(p|\mathbf{y}, \mathbf{w}) = \sum_{i=1}^{m} \{\log g(y_i; p)\} + \sum_{i=m+1}^{N} \{\log g(w_i; p)\}$$

Because this complete data likelihood is in the exponential family, we can use the simple characterization of the EM algorithm by Dempster, Laird, and Rubin (1977). Suppose that $p^{(v)}$ denotes the current value of p after v cycles of the algorithm. The E-step is the estimation of the complete data sufficient statistic $t(\mathbf{x}) = \sum_{i=1}^{N} x_i$ by finding

$$\begin{aligned} t^{(v)} &= E_{\mathbf{X}|\mathbf{Y}, p^{(v)}} \left[\sum_{i=1}^{N} x_i \right] \\ &= \sum_{i=1}^{m} y_i + E_{\mathbf{X}|\mathbf{Y}, p^{(v)}} \left[\sum_{i=m+1}^{m+U} W_i \right] \\ &= \sum_{i=1}^{m} y_i + \left[u(p^{(v)}) \cdot E_{\mathbf{W}|p^{(v)}} [W] \right] \end{aligned}$$

$$= \sum_{i=1}^{m} y_i + \frac{m}{F_g(k;p^{(v)})} \cdot \left[\frac{1}{p^{(v)}} - \sum_{j=1}^{k} j \cdot g(j;p^{(v)}) \right]$$

where

$$\begin{aligned} u(p^{(v)}) &= E_{\mathbf{X}|\mathbf{Y},p^{(v)}}[U] \\ &= E_{\mathbf{X}|\mathbf{Y},p^{(v)}}[N] - m \\ &= \frac{m\left(1 - F_g(k;p^{(v)})\right)}{F_g(k;p^{(v)})} \end{aligned}$$

and

$$\begin{aligned} E_{\mathbf{W}|p^{(v)}}[W] &= \sum_{w=k+1}^{\infty} w \cdot \frac{g(w;p^{(v)})}{1 - F_g\left(k;p^{(v)}\right)} \\ &= \left[\sum_{j=1}^{\infty} j \cdot \frac{g(j;p^{(v)})}{1 - F_g\left(k;p^{(v)}\right)} \right] - \left[\sum_{j=1}^{k} j \cdot \frac{g(j;p^{(v)})}{1 - F_g\left(k;p^{(v)}\right)} \right] \\ &= \left[\frac{1}{p^{(v)}} - \sum_{j=1}^{k} j \cdot g(j;p^{(v)}) \right] \frac{1}{1 - F_g\left(k;p^{(v)}\right)}. \end{aligned}$$

The M-step is to compute

$$p^{(v+1)} = \frac{m}{F_g\left(k;p^{(v)}\right)} \cdot \left\{ \sum_{i=1}^{m} y_i + \frac{m}{F_g(k;p^{(v)})} \cdot \left[\frac{1}{p^{(v)}} - \sum_{j=1}^{k} j \cdot g(j;p^{(v)}) \right] \right\}^{-1}.$$

We iterate through the E and M steps until $p^{(v+1)} - p^{(v)}$ is negligible then use (2.8) to estimate $\widehat{N}$.

Geometric Finite Mixture

As in A.1, let $\mathbf{W} \equiv \{W_i = X_i : i = m+1, ..., N\}$ denote the hypothetical times to first capture that are not observable. Now, the complete data are represented by $(\mathbf{y}, \mathbf{w})^T$. The length of $\mathbf{w}$ is the number of unobserved times to first capture. For the finite geometric mixture, the marginal mass function of y is

$$f(y;\theta) = \pi g(y;p_1) + (1-\pi)\, g(y;p_2).$$

Let Z_i be a latent indicator variable such that $Z_i = 1$ if animal i belongs to group 1 and $Z_i = 0$ otherwise. The marginal mass function for z is

$$f(z;\pi) = \pi^z (1-\pi)^{(1-z)}.$$

The conditional mass function for y given z is

$$f\left(y|z;p_1,p_2\right)=\left[g\left(y;p_1\right)\right]^{z}\left[g\left(y;p_2\right)\right]^{(1-z)}.$$

The joint mass function for y and z is

$$\begin{aligned} m\left(y,z;\theta\right) &= f\left(y|z;p_1,p_2\right)f\left(z;\pi\right) \\ &= \left[\pi g\left(y;p_1\right)\right]^{z}\left[\left(1-\pi\right)g\left(y;p_2\right)\right]^{(1-z)}. \end{aligned}$$

The complete data log likelihood is

$$\begin{aligned} L(\theta|\mathbf{y},\mathbf{w},\mathbf{z}) &= \sum_{i=1}^{m}\left\{\begin{array}{l} z_i\left[\log\pi+\log g(y_i;p_1)\right] \\ \quad +(1-z_i)\left[\log\left(1-\pi\right)+\log g(y_i;p_2)\right] \end{array}\right\} \\ &\quad +\sum_{i=m+1}^{N}\left\{\begin{array}{l} z_i\left[\log\pi+\log g(w_i;p_1)\right] \\ \quad +(1-z_i)\left[\log\left(1-\pi\right)+\log g(w_i;p_2)\right] \end{array}\right\}. \end{aligned}$$

For the EM algorithm, suppose that $\theta^{(v)}$ denotes the current value of θ after v cycles of the algorithm. Define

$$\begin{aligned} Q\left(\theta;\theta^{(v)}\right) &\equiv E_{\mathbf{W},\mathbf{Z}|\mathbf{Y},\theta^{(v)}}\left[L\left(\theta|\mathbf{y},\mathbf{W},\mathbf{Z}\right)\right] \\ &= \sum_{y=1}^{k}n_y\left\{\begin{array}{l} \tau(y;\theta)[\log\pi+\log g(y;p_1)] \\ \quad +\left(1-\tau(y;\theta)\right)\left[\log\left(1-\pi\right)+\log g(y;p_2)\right] \end{array}\right\} \\ &\quad +u\left(\theta^{(v)}\right)E_{W|\theta^{(v)}}\left[\begin{array}{l} \tau(W;\theta^{(v)})\left[\log\pi+\log g(W;p_1)\right] \\ \quad +\left(1-\tau(W;\theta^{(v)})\right)\left[\log\left(1-\pi\right)+\log g(W;p_2)\right] \end{array}\right] \end{aligned}$$

where

$$u\left(\theta^{(v)}\right)=\frac{m\left(1-F_2(k;\theta^{(v)})\right)}{F_2(k;\theta^{(v)})}$$

and

$$\tau(x_i;\theta)\equiv P\left[Z_i=1|x_i\right]=\frac{\pi g(x_i;p_1)}{f_2(x_i;\theta)}.$$

Now, the EM algorithm is performed in the context of the general form described by Dempster et al. (1977) and McLachlan and Jones (1988). For the E-step, compute $Q\left(\theta;\theta^{(v)}\right)$. Then in the M-step, maximize $Q\left(\theta;\theta^{(v)}\right)$ with respect to θ. The MLE's of the M-step are

$$\pi^{(v+1)}=\frac{u(\theta^{(v)})E_{W|\theta^{(v)}}\left[\tau(W;\theta^{(v)})\right]+\sum_{y=1}^{k}n_y\tau(y;\theta^{(v)})}{u(\theta^{(v)})+m},$$

$$p_1^{(v+1)} = \frac{u(\theta^{(v)})E_{W|\theta^{(v)}}\left[\tau(W;\theta^{(v)})\right] + \sum_{y=1}^{k} n_y \tau(y;\theta^{(v)})}{u(\theta^{(v)})E_{W|\theta^{(v)}}\left[W\tau(W;\theta^{(v)})\right] + \sum_{y=1}^{k} y n_y \tau(y;\theta^{(v)})},$$

and

$$p_2^{(v+1)} = \frac{u(\theta^{(v)})\left\{1 - E_{W|\theta^{(v)}}\left[\tau(W;\theta^{(v)})\right]\right\} + \sum_{y=1}^{k} n_y \left(1 - \tau(y;\theta^{(v)})\right)}{u(\theta^{(v)})\left\{E_{W|\theta^{(v)}}\left[W\right] - E_{W|\theta^{(v)}}\left[W\tau(W;\theta^{(v)})\right]\right\} + \sum_{y=1}^{k} y n_y \tau(y;\theta^{(v)})},$$

where the expectations

$$E_{W|\theta}\left[\tau(W;\theta)\right] = \frac{\pi}{1 - F_2(k;\theta)}\left(1 - \sum_{y=1}^{k} g(y;p_1)\right),$$

$$E_{W|\theta}\left[W\tau(W;\theta)\right] = \frac{\pi}{1 - F_2(k;\theta)}\left[\frac{1}{p_1} - \sum_{y=1}^{k} y g(y;p_1)\right],$$

and

$$E_{W|\theta}\left[W\right] = \frac{1}{1 - F_2(k;\theta)}\left[\frac{\pi}{p_1} + \frac{(1-\pi)}{p_2} - \sum_{y=1}^{k} y f_2(y;\theta)\right]$$

come from the E-step. As before, we iterate through the E and M steps until $\theta^{(v+1)} - \theta^{(v)}$ is negligible then use (2.8) to estimate N.

APPENDIX B. DERIVATIVES FOR REMOVAL MODELS

These derivatives are needed for calculating the inverse observed information matrix. As in Section 2, define the pmf for the geometric distribution as

$$g(x) = p(1-p)^{x-1}, \qquad x = 1, 2, ...$$

This can be written as

$$\begin{aligned} g(1) &= p \text{ and} \\ g(x) &= g(x-1)(1-p), \quad x = 2, ... \end{aligned}$$

The first derivative is

$$\frac{\partial g(1)}{\partial p} = 1$$

and

$$\frac{\partial g(x)}{\partial p} = \left[\frac{\partial g(x-1)}{\partial p}(1-p)\right] - g(x-1), \qquad x = 2, ...$$

The second derivative is

$$\frac{\partial^2 g(1)}{\partial p \partial p} = 0$$

and

$$\frac{\partial^2 g(x)}{\partial p \partial p} = \left[\frac{\partial^2 g(x-1)}{\partial p \partial p}(1-p)\right] - 2\frac{\partial g(x-1)}{\partial p}, \qquad x = 2, ...$$

For $x > 1$, these are calculated recursively (i.e., each function at x depends on the function at $x-1$).

For the geometric finite mixture, the log likelihood (2.5) can be expressed as

$$L = -M \log F(k) + \sum_{y=1}^{k} m_y \log f(y; \theta).$$

Define $A = -M \log F(k)$ and $B = \sum_{y=1}^{k} m_y \log f(y;\theta)$, where $\theta = (\pi \quad p_1 \quad p_2)'$. The first partial derivative is

$$\begin{aligned}
\frac{\partial L}{\partial \theta_i} &= \frac{\partial A}{\partial \theta_i} + \frac{\partial B}{\partial \theta_i}, \\
\frac{\partial A}{\partial \theta_i} &= \frac{-M}{F(k)} \frac{\partial F(k)}{\partial \theta_i}, \\
\frac{\partial B}{\partial \theta_i} &= \sum_{y=1}^{k} \left[\frac{m_y}{f(y;\theta)} \frac{\partial f(y;\theta)}{\partial \theta_i} \right], \qquad i = 1,2,3.
\end{aligned}$$

The second partial derivative is

$$\begin{aligned}
\frac{\partial^2 L}{\partial \theta_i \partial \theta_j} &= \frac{\partial^2 A}{\partial \theta_i \partial \theta_j} + \frac{\partial^2 B}{\partial \theta_i \partial \theta_j}, \\
\frac{\partial^2 A}{\partial \theta_i \partial \theta_j} &= \left[\frac{M}{F(k)^2} \frac{\partial F(k)}{\partial \theta_i} \frac{\partial F(k)}{\partial \theta_j} \right] - \left[\frac{M}{F(k)} \frac{\partial^2 F(k)}{\partial \theta_i \partial \theta_j} \right], \\
\frac{\partial^2 B}{\partial \theta_i \partial \theta_j} &= \sum_{y=1}^{k} \left\{ \left[\frac{-m_y}{f(y;\theta)^2} \frac{\partial f(y;\theta)}{\partial \theta_i} \frac{\partial f(y;\theta)}{\partial \theta_j} \right] + \left[\frac{m_y}{f(y;\theta)} \frac{\partial^2 f(y;\theta)}{\partial \theta_i \partial \theta_j} \right] \right\}
\end{aligned}$$

for i and $j = 1,2,3$. To simplify notation, define $f = f(y;\theta)$, $g_1 = g(y;p_1)$, and $g_2 = g(y;p_2)$. Elements in the above formulas are

$$\begin{aligned}
\frac{\partial f}{\partial \pi} &= g_1 - g_2, \\
\frac{\partial f}{\partial p_1} &= \pi \frac{\partial g_1}{\partial p_1}, \\
\frac{\partial f}{\partial p_2} &= (1-\pi) \frac{\partial g_2}{\partial p_2},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial p_1 \partial p_1} &= \pi \frac{\partial^2 g_1}{\partial p_1 \partial p_1}, \\
\frac{\partial^2 f}{\partial p_2 \partial p_2} &= (1-\pi) \frac{\partial^2 g_2}{\partial p_2 \partial p_2}, \\
\frac{\partial^2 f}{\partial \pi \partial p_1} &= \frac{\partial g_1}{\partial p_1}, \\
\frac{\partial^2 f}{\partial \pi \partial p_2} &= -\frac{\partial g_2}{\partial p_2}, \\
\frac{\partial^2 f}{\partial p_1 \partial p_2} &= 0,
\end{aligned}$$

$$\frac{\partial F(k)}{\partial \pi} = \sum_{y=1}^{k} (g_1 - g_2),$$

$$\frac{\partial F(k)}{\partial p_1} = \sum_{y=1}^{k} \pi \frac{\partial g_1}{\partial p_1},$$
$$\frac{\partial F(k)}{\partial p_2} = \sum_{y=1}^{k} (1-\pi) \frac{\partial g_2}{\partial p_2},$$

$$\frac{\partial^2 F(k)}{\partial p_1 \partial p_1} = \sum_{y=1}^{k} \pi \frac{\partial^2 g_1}{\partial p_1 \partial p_1},$$
$$\frac{\partial^2 F(k)}{\partial p_2 \partial p_2} = \sum_{y=1}^{k} (1-\pi) \frac{\partial^2 g_2}{\partial p_2 \partial p_2},$$
$$\frac{\partial^2 F(k)}{\partial \pi \partial p_1} = \sum_{y=1}^{k} \frac{\partial g_1}{\partial p_1},$$
$$\frac{\partial^2 F(k)}{\partial \pi \partial p_2} = -\sum_{y=1}^{k} \frac{\partial g_2}{\partial p_2}, \text{ and}$$
$$\frac{\partial^2 F(k)}{\partial p_1 \partial p_2} = 0.$$

APPENDIX C. COMPARISON TO JACKKNIFE ESTIMATORS

Table C1.
Exact Bias and RMSE for Pollock and Otto (1983) jackknife on 2-group heterogeneous populations.

p_1	p_2	k	π	bias	RMSE
0.2	0.9	5	0.9	0.007	17.17
			0.7	0.021	15.15
			0.5	0.035	12.81
		10	0.9	24.159	31.49
			0.7	18.790	25.90
			0.5	13.422	20.18
0.1	0.9	5	0.9	-59.042	61.35
			0.7	-45.906	48.23
			0.5	-32.770	35.11
		10	0.9	0	25.05
			0.7	0	22.09
			0.5	0	18.67
0.2	0.5	5	0.9	1.875	17.58
			0.7	5.625	17.11
			0.5	9.375	17.45
		10	0.9	24.115	31.63
			0.7	19.259	26.33
			0.5	14.203	20.90
0.1	0.5	5	0.9	-57.174	59.65
			0.7	-40.302	43.33
			0.5	-23.430	27.63
		10	0.9	0.156	25.08
			0.7	0.469	22.21
			0.5	0.781	18.90

Table C2.

Simulation results for removal of 2-group heterogeneous populations.

				jb*			P2**		
p_1	p_2	k	π	n	bias	RMSE	n	bias	RMSE
0.2	0.9	5	0.9	500	-0.6	17.6	368	14.8	101.4
			0.7	1400	0.1	15.1	976	20.6	83.6
			0.5	500	-.32	13.1	379	0.1	79.5
		10	0.9	500	23.0	31.0	468	4.6	38.8
			0.7	1400	18.4	25.9	1395	4.3	35.3
			0.5	500	13.4	20.2	496	3.9	18.8
0.1	0.9	5	0.9	500	-57.7	60.1	294	-15.6	133.1
			0.7	500	-44.7	47.0	323	-27.1	135.7
			0.5	500	-32.1	34.5	328	-38.9	89.2
		10	0.9	500	0.6	24.6	434	35.6	127.9
			0.7	500	0.1	21.6	462	25.8	101.8
			0.5	500	0.7	18.7	454	17.5	75.0
0.2	0.5	5	0.9	500	0.7	16.8	390	-7.3	78.3
			0.7	1400	5.4	16.7	1249	-55.9	114.7
			0.5	500	8.8	17.1	380	-7.1	101.5
		10	0.9	500	23.6	31.1	500	-8.1	56.3
			0.7	1400	19.6	26.7	1272	-45.7	102.8
			0.5	500	14.8	21.3	368	-16.6	70.4
0.1	0.5	5	0.9	500	-57.3	59.9	339	-41.0	112.3
			0.7	500	-40.6	43.8	357	-55.0	98.0
			0.5	500	-23.2	27.8	358	-50.7	84.5
		10	0.9	500	0.9	24.6	360	3.9	107.5
			0.7	500	1.0	22.0	347	-14.6	92.6
			0.5	500	1.4	19.4	349	-24.0	82.2

* Pollock and Otto (1983) jackknife; ** 2-group finite mixture;

Table C3.

Simulation results for mark-recapture of 2-group heterogeneous populations.

				jack*			PeqC2**		
p_1	p_2	k	π	n	bias	RMSE	n	bias	RMSE
0.2	0.9	5	0.9	500	41.2	46.6	369	18.6	71.4
			0.7	1400	32.2	37.9	1400	3.2	17.3
			0.5	500	22.1	27.8	500	3.6	15.6
		10	0.9	500	24.7	26.5	469	6.6	26.1
			0.7	1400	19.3	21.1	1400	0.2	5.0
			0.5	500	14.0	15.8	500	0.3	4.2
0.1	0.9	5	0.9	500	-2.6	19.2	500	17.4	61.5
			0.7	500	-1.6	17.2	499	18.2	64.5
			0.5	500	-1.1	14.9	499	21.8	80.1
		10	0.9	500	41.5	51.5	500	3.2	16.8
			0.7	500	29.8	39.3	500	2.6	14.9
			0.5	500	20.1	28.7	500	2.2	12.7
0.2	0.5	5	0.9	500	39.6	45.5	421	8.9	62.5
			0.7	1400	27.7	34.0	1307	-16.0	94.4
			0.5	500	18.0	24.3	396	-1.0	84.3
		10	0.9	500	25.4	27.2	498	2.5	10.5
			0.7	1400	20.0	21.8	1400	1.3	6.7
			0.5	500	14.8	16.3	500	1.3	5.9
0.1	0.5	5	0.9	500	-4.8	19.9	411	17.5	125.2
			0.7	500	-7.8	22.1	381	-9.0	111.9
			0.5	500	-11.1	20.8	356	-20.3	97.8
		10	0.9	500	40.5	49.8	500	5.1	20.6
			0.7	500	30.4	40.3	500	4.7	20.0
			0.5	500	19.8	30.5	500	5.8	20.4

* Burnham and Overton (1978) jackknife; ** 2-group finite mixture ($\mathbf{p} = \mathbf{c}$).

APPENDIX D. SIMULATIONS OF HOMOGENEOUS POPULATIONS

Table D1

Simulation of homogeneous removal with $p = 0.1$, $N = 200$, *and 1000 replicates.*

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	P	859	30.4	144.7	147.9	83	87.9
	P2eq1	824	4.8	154.8	154.9	46	28.5
	P2	440	9.8	125.2	125.6	-	325.8
	AIC	815	34.2	149.0	152.8	83	88.5
	jb	1000	-65.0	17.4	67.3	7	16.1
	Zippin	895	12.1	126.8	127.4	79	72.4
10	P	1000	11.3	46.1	47.5	89	33.5
	P2eq1	998	27.4	75.8	80.6	90.1	36.4
	P2	692	25.9	91.7	95.3	99	103.4
	AIC	998	18.3	60.9	63.6	90	35.0
	jb	1000	-0.9	26.7	26.7	91	25.1
	Zippin	1000	4.1	40.7	40.9	88	31.9

* P: homogeneous geometric; P2eq1: 2-group geometric ($p_2 = 1$); P2: 2-group geometric; AIC: estmate chosen by AIC among P, P2eq1, and P2; jb: Pollock and Otto (1983) jackknife; Zippin: Zippin (1956).

Table D2

Simulation of homogeneous mark-recapture with $p = c = 0.1$, $N = 200$, and 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	PC2	390	76.0	167.2	183.7	93	126.4
	PeqC2	670	21.9	86.6	89.4	98	53.1
	jack	1000	-4.3	19.0	19.5	94	20.0
	null	1000	6.8	41.5	42.0	95	37.3
10	PC2	534	42.3	124.6	131.6	97	68.9
	PeqC2	776	12.9	46.9	48.7	97	30.3
	jack	1000	45.1	31.4	55.0	50	18.8
	null	1000	0.06	16.0	16.0	96	16.0

* PC2: 2-group finite mixture; PeqC2: 2-group finite mixture ($\mathbf{p = c}$); jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

Table D3

Simulation of homogeneous removal with $p = 0.2$, $N = 200$, and 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	P	1000	12.7	48.4	50.1	90	30.2
	P2eq1	985	27.0	95.5	99.2	86	36.2
	P2	689	20.5	61.4	64.8	98	123.3
	AIC	988	21.0	67.42	70.6	91	32.0
	jb	1000	0.05	18.2	18.2	94	17.9
	Zippin	1000	5.9	42.5	42.9	90	29.0
10	P	1000	1.0	9.2	9.3	88	6.9
	P2eq1	1000	2.6	10.2	10.5	93	8.0
	P2	766	7.7	31.8	32.7	99	20.7
	AIC	999	1.9	10.5	10.7	88	7.2
	jb	1000	26.7	21.0	34.0	88	21.2
	Zippin	1000	-0.6	8.9	9.0	92	8.0

* P: homogeneous geometric; P2eq1: 2-group geometric ($p_2 = 1$); P2: 2-group geometric; AIC: estmate chosen by AIC among P, P2eq1, and P2; jb: Pollock and Otto (1983) jackknife; Zippin: Zippin (1956).

Table D4

Simulation of homogeneous mark-recapture with $p = 0.2$, $N = 200$, and 1000 replicates.

k	method	n	bias	sample sd	sample RMSE	% coverage	median se
5	PC2	537	32.1	82.1	88.2	98	76.2
	PeqC2	756	10.7	38.6	40.1	99	30.1
	jack	1000	47.9	23.0	53.2	37	21.8
	null	1000	1.0	15.8	15.8	95	15.7
10	PC2	682	9.4	24.1	25.8	97	13.6
	PeqC2	866	4.1	15.3	15.9	92	7.7
	jack	1000	27.1	9.1	28.6	13	9.5
	null	1000	-0.4	5.8	5.8	95	5.8

* PC2: 2-group finite mixture; PeqC2: 2-group finite mixture ($\mathbf{p} = \mathbf{c}$); jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

Table D5

Simulation of homogeneous removal with $p = 0.355$, $N = 722$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	P	1000	1.1	17.8	17.8	89	14.1
	P2eq1	1000	6.4	20.9	21.8	97	19.6
	P2	748	16.7	83.4	85.1	100	46.4
	AIC	999	4.3	26.8	27.2	90	14.4
	jb	1000	96.5	27.6	100.4	5	29.7
	Zippin	1000	-0.5	17.6	17.6	94	16.8
10	P	1000	0.2	3.5	3.5	67	1.7
	P2eq1	1000	0.7	3.6	3.7	76	2.2
	P2	792	2.6	10.4	10.7	90	4.6
	AIC	1000	0.7	4.6	4.6	69	1.8
	jb	1000	36.2	20.8	41.8	69	21.2
	Zippin	1000	-0.4	3.5	3.5	95	3.4
20	P	1000	-0.1	0.4	0.4	0	0.04
	P2eq1	1000	-0.1	0.4	0.4	0	0.05
	P2	914	0.4	4.5	4.5	74	0.2
	AIC	1000	0.1	4.1	4.1	2	0.04
	jb	1000	-0.2	0.4	0.5	84	0
	Zippin	1000	-0.2	0.4	0.5	15	0.3

* P: homogeneous geometric; P2eq1: 2-group geometric ($p_2 = 1$); P2: 2-group geometric; AIC: estmate chosen by AIC among P, P2eq1, and P2; jb: Pollock and Otto (1983) jackknife; Zippin: Zippin (1956).

APPENDIX E. SIMULATIONS OF HETEROGENEOUS POPULATIONS

Table E1

Parameters for Simulations

name	N	group	π	p	c
A	200	1	0.7	0.2	0.2
		2	0.3	0.9	0.9
B	100	1	0.333	0.1	0.1
		2	0.333	0.2	0.2
		3	0.333	0.5	0.5
Taxi	420	1	0.978	0.104	0.132
		2	0.022	1	0.267
Vole	133	1	0.704	0.117	0.493
		2	0.158	0.883	0.775
		3	0.138	1	0.223

Table E2

Removal Simulation A, $N = 200$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	P	1000	-35.7	7.4	36.5	1	3.4
	P2eq1	997	-3.6	48.0	48.1	73	18.1
	P2	691	16.6	70.6	72.5	90	52.1
	AIC	995	-3.5	48.2	48.3	72	18.1
	jb	1000	0.1	15.0	15.0	93	14.8
	Zippin	1000	-36.8	7.3	37.5	1	4.4
10	P	1000	-10.7	4.1	11.5	5	1.5
	P2eq1	1000	-1.9	7.3	7.5	79	5.2
	P2	980	6.9	26.5	27.4	91	9.5
	AIC	998	4.7	26.4	26.9	82	6.1
	jb	1000	18.5	18.1	25.8	92	19.0
	Zippin	1000	-11.5	4.1	12.2	11	2.4
20	P	1000	-1.6	1.5	2.2	6	0.2
	P2eq1	1000	-0.6	0.6	1.6	1.7	0.6
	P2	1000	-0.2	1.8	1.8	76	0.9
	AIC	1000	-0.4	0.9	1.8	64	0.7
	jb	1000	-2.0	1.5	2.5	15	0
	Zippin	1000	-2.0	1.5	2.5	56	0.6

* P: homogeneous geometric; P2eq1: 2-group geometric ($p_2 = 1$); P2: 2-group geometric; AIC: estmate chosen by AIC among P, P2eq1, and P2; jb: Pollock and Otto (1983) jackknife; Zippin: Zippin (1956).

Table E3

Mark-recapture Simulation A, $N = 200$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	PC2	993	14.4	66.6	68.2	89	24.8
	PeqC2	1000	3.3	17.4	17.7	92	13.0
	jack	1000	32.5	20.2	38.3	53	17.3
	null	1000	-42.1	6.0	42.6	0	2.2
10	PC2	1000	0.9	7.4	7.5	90	5.9
	PeqC2	1000	0.3	5.0	5.0	75	2.8
	jack	1000	19.3	8.3	21.0	30	8.0
	null	1000	-15.0	3.6	15.4	0	0.7
20	PC2	1000	-0.4	0.7	1.6	1.7	0.7
	PeqC2	1000	-0.4	0.3	1.5	35	0.3
	jack	1000	5.6	6.0	8.2	84	3.6
	null	1000	-2.0	1.5	2.5	15	0.07

* PC2: 2-group finite mixture; PeqC2: 2-group finite mixture ($\mathbf{p} = \mathbf{c}$); jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

Table E4

Removal Simulation B, $N = 100$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	P	1000	-18.7	9.8	21.1	29	5.9
	P2eq1	991	-11.4	22.5	25.2	57	10.4
	P2	646	-3.2	34.8	35.0	90	51.4
	AIC	990	-14.3	20.5	25.0	38	6.5
	jb	1000	-7.6	10.6	13.1	82	11.0
	Zippin	1000	-20.7	8.8	22.5	26	6.1
10	P	1000	-10.5	4.4	11.4	13	1.9
	P2eq1	1000	-8.4	5.9	10.3	37	3.2
	P2	626	1.6	29.7	29.7	80	10.0
	AIC	990	-6.4	21.2	22.1	30	2.5
	jb	1000	5.4	13.7	14.7	90	13.4
	Zippin	1000	-11.5	4.2	12.2	17	2.6
20	P	1000	-4.0	2.1	4.6	7	0.4
	P2eq1	1000	-3.6	2.3	4.2	19	0.7
	P2	937	0.1	9.0	9.0	71	2.6
	AIC	999	-0.9	8.8	8.8	44	0.9
	jb	1000	-4.8	2.0	5.2	0	0
	Zippin	1000	-46	2.1	5.1	19	0.8

* P: homogeneous geometric; P2eq1: 2-group geometric ($p_2 = 1$); P2: 2-group geometric; AIC: estmate chosen by AIC among P, P2eq1, and P2; jb: Pollock and Otto (1983) jackknife; Zippin: Zippin (1956).

Table E5

Mark-recapture Simulation B, $N = 100$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	PC2	677	0.2	41.6	41.6	73	18.6
	PeqC2	670	4.7	43.0	43.2	74	17.9
	jack	1000	2.4	13.6	13.8	81	9.8
	null	1000	-22.2	5.8	22.96	3	4.0
10	PC2	980	3.6	27.1	27.4	79	8.9
	PeqC2	997	0.6	13.0	13.0	82	5.6
	jack	1000	8.2	12.5	15.0	87	6.1
	null	1000	-13.3	3.6	13.7	0	1.6
20	PC2	1000	-1.9	2.9	3.5	60	1.5
	PeqC2	1000	-2.1	2.4	3.2	42	0.8
	jack	1000	5.2	4.5	6.8	90	4.3
	null	1000	-4.8	2.0	5.2	0	0.4

* PC2: 2-group finite mixture; PeqC2: 2-group finite mixture ($\mathbf{p} = \mathbf{c}$); jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

Table E6

Taxi removal simulation, $N = 420$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	P	995	-46.1	139.7	147.1	65	74.7
	P2eq1	893	63.7	282.3	289.4	86	170
	P2	617	68.2	268.2	276.7	99	518
	AIC	923	14.1	249.6	250.0	71	89
	jb	1000	-127.5	24.3	129.9	0	23.7
	Zippin	997	-59.8	130.8	143.8	62	70.3
10	P	1000	-18.4	39.3	43.3	80	33.9
	P2eq1	1000	13.7	65.8	67.2	91	52.7
	P2	857	45.2	132.0	139.5	97	89.8
	AIC	989	7.6	83.0	83.3	86	41.4
	jb	1000	5.7	36.8	37.2	95	37.9
	Zippin	1000	-22.8	37.8	44.2	80	34.8
20	P	1000	-15.3	10.9	18.8	51	7.9
	P2eq1	1000	-11.7	12.2	16.9	66	9.3
	P2	977	-9.9	13.9	17.1	80	11.4
	AIC	1000	-12.6	13.0	18.1	60	8.5
	jb	1000	-50.9	6.9	51.3	0	0
	Zippin	1000	-16.6	10.8	19.8	62	9.8

* P: homogeneous geometric; P2eq1: 2-group geometric ($p_2 = 1$); P2: 2-group geometric; AIC: estmate chosen by AIC among P, P2eq1, and P2; jb: Pollock and Otto (1983) jackknife; Zippin: Zippin (1956).

Table E7

Taxi mark-recapture simulation, $N = 420$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	PC2	624	87.4	271.1	284.9	96	360
	PeqC2	888	-64.4	59.0	87.3	69	48.7
	jack	1000	-23.7	28.6	37.2	84	28.5
	null	1000	-80.1	33.9	87.0	34	31.5
10	PC2	921	50.3	157.9	165.7	95	68.4
	PeqC2	903	-35.2	43.7	56.1	50	17.4
	jack	1000	37.3	32.5	49.5	73	37.3
	null	1000	-51.6	16.7	54.2	12	14.6
20	PC2	989	-10.1	12.9	16.4	72	9.8
	PeqC2	975	-26.4	8.6	27.8	9	3.6
	jack	1000	20.0	10.9	22.8	63	11.7
	null	1000	-28.3	7.7	29.4	1	5.5

* PC2: 2-group finite mixture; PeqC2: 2-group finite mixture ($\mathbf{p} = \mathbf{c}$); jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

Table E8

Vole removal simulation, $N = 133$, 1000 replicates

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	P	1000	-45.8	6.4	46.3	0	2.2
	P2eq1	867	-2.5	72.1	72.1	63	25.4
	P2	562	3.9	78.5	78.6	79	52.4
	AIC	887	-3.8	71.7	71.8	61	23.9
	jb	1000	-24.0	12.4	27.0	43	11.0
	Zippin	1000	-46.8	6.2	47.3	0	2.7
10	P	1000	-22.3	5.2	23.0	0	2.0
	P2eq1	995	1.5	27.2	27.2	82	14.7
	P2	911	18.5	55.6	58.6	92	26.0
	AIC	988	8.9	43.1	44.0	84	16.5
	jb	1000	5.1	17.4	18.1	93	19.0
	Zippin	1000	-23.3	5.2	23.8	1	2.9
20	P	1000	-7.1	3.1	7.7	5	0.7
	P2eq1	1000	-2.2	4.7	5.2	71	2.9
	P2	1000	-0.9	5.8	5.8	80	3.8
	AIC	1000	-1.5	5.5	5.7	75	3.3
	jb	1000	-8.8	2.8	9.2	0	0
	Zippin	1000	-7.7	3.1	8.3	12	1.4

* P: homogeneous geometric; P2eq1: 2-group geometric ($p_2 = 1$); P2: 2-group geometric; AIC: estmate chosen by AIC among P, P2eq1, and P2; jb: Pollock and Otto (1983) jackknife; Zippin: Zippin (1956).

Table E9

Vole mark-recapture simulation, $N = 133$, 1000 replicates.

k	method*	n	bias	sample sd	sample RMSE	% coverage	median se
5	PC2	868	-31.1	32.0	44.6	32	8.7
	PeqC2	987	-39.3	14.0	41.7	9	4.2
	PC3	447	12.8	87	88	50	13.3
	jack	1000	-32.6	8.3	33.6	5	5.5
	null	1000	-48.2	5.2	48.5	0	1.8
10	PC2	1000	-15.0	9.8	17.9	37	5.7
	PeqC2	1000	-25.3	4.4	25.7	0	0.6
	PC3	909	21.0	56.9	60.7	54	9.2
	jack	1000	-17.6	6.9	18.9	8	4.1
	null	1000	-27.3	4.3	27.6	0	0.6
20	PC2	1000	-5.0	3.5	6.1	35	1.7
	PeqC2	1000	-8.8	2.8	9.2	0	0.01
	PC3	992	-1.1	5.1	5.2	67	2.4
	jack	1000	-7.0	4.8	8.4	15	1.4
	null	1000	-8.8	2.8	9.3	0	0.04

* PC2: 2-group finite mixture; PeqC2: 2-group finite mixture ($\mathbf{p} = \mathbf{c}$); PC3: 3-group finite mixture; jack: Burnham and Overton (1978) jackknife; null: Otis et. al. (1978) multinomial.

BIBLIOGRAPHY

Bain, L. J. and M. Engelhardt. (1989). *Introduction to Probability and Mathematical Statistics.* PWS-Kent, Boston.

Buckland, S. T., I. B. J. Goodie, and D. L. Borchers. (2000). Wildlife population assessment: past development and future directions. *Biometrics* **56,** 1-12.

Burnham, K. P. (1972). Estimation of population size in multiple capture-recapture studies when capture probabilities vary among animals. unpublished dissertation. Oregon State University.

Burnham, K. P. and W. S. Overton. (1978). Estimation of the size of a closed population when capture probablilities vary among animals. *Biometrika* **65,** 625-633.

Burnham, K. P. and W. S. Overton. (1979). Robust estimation of population size when capture probabilities vary among animals. *Ecology* **60,** 927-936.

Carothers, A. D. (1973). Capture-recapture methods applied to a population with known parameters. *Journal of Animal Ecology* **42,** 125-146.

Cohen, A. C. (1966). A note on certain discrete mixed distributions. *Biometrics* **22,** 566-572.

Dempster, A. P., N. P. Laird, and D. B. Rubin. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statistical Society Series B* **39,** 1-38.

Everitt, B. S. and D. J. Hand. (1981). *Finite Mixture Distributions.* Chapman and Hall, New York.

Getz, L. L., J. E. Hofman, B. J. Klatt, L. Verner, F. R. Cole, and R. L. Lindroth. (1987). Fourteen years of population fluctuations of *Microtus ochrogaster* and *M. pennsylvanicus* in east central Illinois. *Canadian Journal of Zoology* **65,** 1317-1325.

Hansen, P. A. and G. A. Pearse. (1995). Abundance and composition of Northern Pike in Volkmar and Deadman Lakes, 1994. Alaska Department of Fish and Game, Division of Sport Fish, Fishery Data Series No. 95-7.

Hooke, J. and T. A. Jeeves. (1961). Direct search solution to numerical and statistical problems. *Journal of the Association of Computational Machines* **8,** 212-229.

Lindsay, B. G. and K. Roeder. (1992). Residual diagnostics for mixture models. *Journal of the American Statistical Association* **87,** 785-794.

McLachlan, G. J. and P. N. Jones. (1988). Fitting mixture models to grouped and truncated data via the EM algorithm. *Biometrics* **44,** 571-578.

Natarajan, R. and C. E. McCulloch. (1999). Modeling heterogeneity in Nest Survival Data. *Biometrics* **55,** 553-559.

Nelder, J. A. and R. Mead. (1965). A simplex method for function minimization. *Compt. J.* **7,** 308-313.

Norris, J. L. III and K. H. Pollock. (1996). Nonparametric MLE under two closed capture-recapture models with heterogeneity. *Biometrics* **52,** 639-649.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. (1978). Statistical Inference from capture data on closed animal populations. *Wildlife Monographs* **62,** 1-135.

Pollock, K. H. (2000). Capture-recapture models. *Journal of the American Statistical Association* **95,** 293-296.

Pollock, K. H. and M. C. Otto. (1983). Robust estimation of population size in closed animal populations from capture-recapture experiments. *Biometrics* **39,** 1035-1049.

Rexstad, E. and K. P. Burnham. (1991). User's guide for interactive program CAPTURE, Abundance estimation of closed animal populations. Colorado State University, Fort Collins, Colorado, USA.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters.* Edward Arnold, London.

Wu, C. F. J. (1983). On the convergence properties of the EM algorithm. *The Annals of Statistics* **11,** 95-103.

Young, H., J. Neess, and J. T. Emlen, Jr. (1952). Heterogeneity of trap response in a population of house mice. *Journal of Wildlife Management* **16,** 169-180.

Zippin C. (1956). An evaluation of the removal method of estimating animal populations. *Biometrics* **12,** 103-189.